

SCHOOL OF CIVIL ENGINEERING

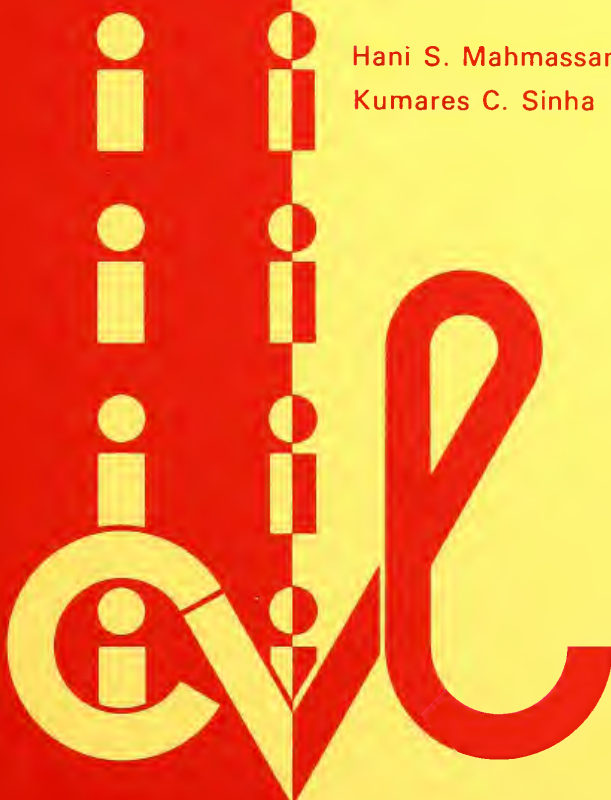


JOINT HIGHWAY RESEARCH PROJECT

FHWA/ISHC/JHRP-78/13

FRAMEWORK FOR THE TRANSFERABILITY
OF TRIP GENERATION PARAMETERS FOR
SMALL URBAN AREAS IN INDIANA

Hani S. Mahmassani
Kumares C. Sinha



PURDUE UNIVERSITY
INDIANA STATE HIGHWAY COMMISSION

Interim Report
FRAMEWORK FOR THE TRANSFERABILITY OF TRIP GENERATION
PARAMETERS FOR SMALL URBAN AREAS IN INDIANA

TO: Harold L. Michael, Director
Joint Highway Research Project

FROM: K. C. Sinha, Research Engineer
Joint Highway Research Project

August 2, 1978

Project: C-36-69L

File: 3-7-12

Attached is the first Interim Report on the HPR Part I Study titled "Use of Synthetic Demand Modelling Techniques in Transportation Planning for Small Urban Areas in Indiana". The title of the Report is "Framework for the Transferability of Trip Generation Parameters of Small Urban Areas" and it has been conducted by Mr. Hani S. Mahmassani, Graduate Instructor in Research, of our staff under the direction of Professor K. C. Sinha, Research Engineer.

The Report here is concerned with development of a realistic framework for transferring trip generation parameters from small urban areas to other small urban areas. A recommended procedure for synthetic trip generation analysis is proposed and will be tested and evaluated in research still in progress.

The Report is forwarded for review, comment and acceptance as partial fulfillment of the objectives of the research, not only to the JHRP Board but to ISHC and FHWA as well.

Respectfully submitted,



K. C. Sinha
Research Engineer

KCS:ms

cc: A. G. Altschaeffl
W. L. Dolch
R. L. Eskew
G. D. Gibson
W. H. Goetz
M. J. Gutzwiller
G. K. Hallock

D. E. Hancher
K. R. Hoover
J. F. McLaughlin
R. F. Marsh
R. D. Miles
P. L. Owens
G. T. Satterly

C. F. Scholer
M. B. Scott
K. C. Sinha
C. A. Venable
L. E. Wood
E. J. Yoder
S. R. Yoder

Interim Report
FRAMEWORK FOR THE TRANSFERABILITY OF TRIP GENERATION
PARAMETERS FOR SMALL URBAN AREAS IN INDIANA

by
Hani S. Mahmassani
Graduate Instructor in Research
and

Kumares C. Sinha
Research Engineer

Joint Highway Research Project
Project No.: C-36-69L
File No.: 3-7-12

Prepared as Part of an Investigation
Conducted by
Joint Highway Research Project
Engineering Experiment Station
Purdue University
in cooperation with the
Indiana State Highway Commission
and the
U.S. Department of Transportation
Federal Highway Administration

The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the Federal Highway Administration.

Purdue University
West Lafayette, Indiana
August 2, 1978



Digitized by the Internet Archive
in 2011 with funding from
LYRASIS members and Sloan Foundation; Indiana Department of Transportation

1. Report No. FHWA/ISHC/JHRP-78/13	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle FRAMEWORK FOR THE TRANSFERABILITY OF TRIP GENERATION PARAMETERS FOR SMALL URBAN AREAS IN INDIANA		5. Report Date August 2, 1978	
		6. Performing Organization Code	
7. Author(s) Hani S. Mahmassani and Kumares C. Sinha		8. Performing Organization Report No. JHRP-78-13	
9. Performing Organization Name and Address Joint Highway Research Project Civil Engineering Building Purdue University W. Lafayette, Indiana 47907		10. Work Unit No.	
		11. Contract or Grant No. HPR-1(16) Part I	
12. Sponsoring Agency Name and Address* Indiana State Highway Commission State Office Building 100 North Senate Avenue Indianapolis, Indiana 46204		13. Type of Report and Period Covered Interim Report (Initial)	
		14. Sponsoring Agency Code	
15. Supplementary Notes Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration under a planning study titled "Use of Synthetic Demand Modelling Techniques in Transportation Planning for Small Urban Areas in Indiana"			
16. Abstract Initiating a continuing, comprehensive and coordinated transportation planning process is a considerable task that faces small metropolitan areas as their population reaches 50,000. These areas usually lack the technical know-how as well as the financial capabilities needed to use the battery of computer-oriented planning models more commonly used in larger metropolitan areas. In most cases, such sophistication is not needed because the problems faced by smaller areas are different in nature and in complexity from those in larger areas. This study examined the small and medium sized urban areas (population range: 50,000 to 250,000) of Indiana where full scale transportation studies have already been conducted, in an effort to make more efficient use of the information made available by these studies. This study focused on trip generation (frequency) models and parameters and examined the stability and transferability of these parameters between small and medium sized urban areas at three levels of aggregation: areawide, zonal and household levels. A framework was thus provided for the transferability of trip frequency parameters. Special emphasis was placed on trip generation by category analysis. Synthetic trip generation procedures aimed at eliminating the need for costly home interview O-D surveys were recommended. A methodology based on Bayesian statistics was suggested for updating cross-classified household trip generation rates.			
17. Key Words Transportation Planning; Trip Generation; Disaggregate Analysis; Cross Classification; Bayesian Updating; Zonal Analysis		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 149	22. Price

TABLE OF CONTENTS

	Page
LIST OF TABLES.	v
LIST OF FIGURES	ix
HIGHLIGHT SUMMARY	x
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: GENERAL CHARACTERISTICS OF THE STUDY AREAS	9
CHAPTER 3: DISTRIBUTION OF TRIPS BY PURPOSE	14
Number and Definition of Trip Purposes	16
Variation Between Areas.	17
Use of Statewide Averages.	23
Distribution of Person Trips by Purpose.	24
Variation Within Urban Areas	25
CHAPTER 4: AREAWIDE TRIP FREQUENCY.	32
Factors Affecting Areawide Travel Frequency.	39
Synthetic Forecasting of Areawide Travel Frequency	40
Comparative Analysis of the Household Auto Ownership Distribution.	44
Implications for Trip Frequency Transferability and Synthetic Application	50
CHAPTER 5: TRIP GENERATION ANALYSIS AT THE ZONAL LEVEL.	53
Regression Technique	53
Review of Past Synthetic Efforts Using Zonal Regression Equations.	54
Internal Vehicle Trip Generation Equations Developed in Study Areas.	58
Transferability of the Internal Vehicle Trip Production Equations.	60
Transferability of the Internal Vehicle Trip Attraction Equations.	71
Critical Review of Aggregate Zonal Trip Generation.	77

	Page
CHAPTER 6: TRIP GENERATION ANALYSIS AT THE DISAGGREGATE HOUSEHOLD LEVEL.	79
Disaggregate Household Trip Generation Models.	80
Cross-Classification Procedure	81
Levels of the Independent Variables.	83
Cross-Classified Vehicle Trip Rate Tables for Lafayette and Evansville.	97
Temporal and Spatial Stability of Household Trip Generation Models.	113
Comparison of Cross-Classified Household Trip Rates for Lafayette and Evansville.	115
CHAPTER 7: BAYESIAN UPDATING IN TRANSPORTATION PLANNING	121
Updating Procedures.	122
Use of Bayesian Updating in Transportation	122
Bayesian Updating of Cross-Classified Trip Generation Rates.	123
Small Disaggregate Samples	124
Subjective Judgement as a Source of New Information	129
CHAPTER 8: CONCLUSION	135
Framework for the Transferability of Trip Frequency Parameters.	137
Recommended Procedure for Synthetic Trip Generation Analysis in Small Urban Areas	139
Recommendations for Further Research	140
LIST OF REFERENCES.	142
APPENDIX.	147

LIST OF TABLES

Table	Page
1. Study Areas General Characteristics.	12
2. Number and Percentage of Vehicle Trips by Type and Urban Area	15
3. Distribution of Vehicle Trips by Purpose	18
4. Summary of Chi-Square Test -- Effect of Urban Area on Trip Purpose Mix	22
5. Chi-Square Test of Similarity of Distribution of Trips by Purpose Between Lafayette and Terre Haute	22
6. Distribution of Person Trips by Purpose.	25
7. Distribution of Household Vehicle Trips by Purpose and Household Auto Ownership - Evansville.	29
8. Contingency Table and Summary of Chi-Square Test of Independence of Trip Purpose Distribution and Auto Ownership - Evansville	29
9. Areawide Internal Vehicle Trip Rates: Daily Trips per Household.	34
10. Areawide Internal Vehicle Trip Rates: Daily Trips per Capita	35
11. Areawide Internal Vehicle Trip Rates: Daily Trips per Automobile	36
12. Areawide Internal Person Trip Rates: Daily Trips per Household.	36
13. Areawide Internal Person Trip Rates: Daily Trips per Capita	38
14. Areawide Internal Person Trip Rates: Daily Trips per Automobile	38

Table	Page
15. Relationship Between Average Household Person Trip Frequency and Average Auto Ownership per Household Stratified by Urban Size and Urban Structure	42
16. Comparison of Areawide Household Person Trip Frequency Rates: Predicted vs. Observed.	43
17. Household Distribution by Auto Ownership	45
18. Summary of Chi-Square Test for Household Distribution by Auto Ownership.	47
19. Summary of Chi-Square Tests for Pairwise Comparisons of Household Distribution by Auto Ownership	49
20. Internal Vehicle Trip Production Equations Developed in Study Areas - HBW and HBO	59
21. Non-Home Based Internal Vehicle Trip Production and Attraction Equations Developed in Study Areas.	61
22. Internal Vehicle Trip Attraction Equations Developed in Study Areas - HBW and HBO	62
23. Internal Person Trip Generation Equations - Evansville	63
24. Comparison of Estimated Mean Zonal Trips to Observed Mean Zonal Trips - Home Based Work Productions.	65
25. Summary of the Predictive Ability of HBW Vehicle Trip Production Equations Developed in Study Areas	66
26. Comparison of Estimated Mean Zonal Trips to Observed Mean Zonal Trips - Home Based Other Productions	68
27. Summary of the Predictive Ability of HBO Vehicle Trip Production Equations Developed in Study Areas	69
28. Comparison of Estimated Mean Zonal Trips to Observed Mean Zonal Trips - Non-Home Based Productions.	70
29. Summary of the Predictive Ability of NHB Vehicle Trip Production Equations Developed in Study Areas.	72
30. Comparison of Estimated Mean Zonal Trips to Observed Mean Zonal Trips - Home Based Work Attractions	74
31. Comparison of Estimated Mean Zonal Trips to Observed Mean Zonal Trips - Home Based Other Attractions.	75

Table	Page
32. Comparison of Estimated Mean Zonal Trips to Observed Mean Zonal Trips - Non-Home Based Attractions.	76
33. Cross-Classification of Total Vehicle Trip Rates, LAFAYETTE: 5 Levels of Auto Ownership and 6 Levels of Household Size.	87
34. Cross-Classification of Total Vehicle Trip Rates, LAFAYETTE: 4 Levels of Auto Ownership and 6 Levels of Household Size.	89
35. Sample Test Calculation: Comparison of the Means of Cells (Cars = 1, Size = 6) and (Cars = 1, Size = 5).	91
36. Summary of t-tests for Step 3.	92
37. Cross-Classification of Total Vehicle Trip Rates, LAFAYETTE: 4 Levels of Auto Ownership and 5 Levels of Household Size.	93
38. Summary of t-tests for Step 4.	94
39. Summary of t-tests for Step 5.	95
40. Cross-Classification of HBW Household Vehicle Trip Rates, LAFAYETTE.	98
41. Cross-Classification of HBO Household Vehicle Trip Rates, LAFAYETTE.	99
42. Cross-Classification of NHB Household Vehicle Trip Rates, LAFAYETTE.	100
43. Cross-Classification of HBW Household Vehicle Trip Rates, EVANSVILLE	101
44. Cross-Classification of HBO Household Vehicle Trip Rates, EVANSVILLE	102
45. Cross-Classification of NHB Household Vehicle Trip Rates, EVANSVILLE	103
46. Cross-Classification of ALL Household Vehicle Trip Rates, EVANSVILLE	104
47. T-test for Comparison of Cell Means.	116
48. Cell-by-Cell Comparison of Lafayette and Evansville HBW Trip Rate Tables: Summary of t-tests Results	117

Table	Page
49. Cell-by-Cell Comparison of Lafayette and Evansville HBO Trip Rate Tables: Summary of t-tests Results	118
50. Cell-by-Cell Comparison of Lafayette and Evansville NHB Trip Rate Tables: Summary of t-tests Results	119
51. Summary of Bayesian Updating Techniques for Trip Generation Rates	133
Appendix	
Table	
A1. General Contingency Table, a.b	148

LIST OF FIGURES

Figure	Page
1. Standard Metropolitan Statistical Areas - Indiana.	10
2. Evansville: Purpose Distribution of Vehicle Trips by Auto Ownership.	28
3. Home Based Work Vehicle Trips per Household by Household Size and Auto Ownership -- Lafayette	105
4. Home Based Other Vehicle Trips per Household by Household Size and Auto Ownership -- Lafayette	106
5. Non-Home Based Vehicle Trips per Household by Household Size and Auto Ownership -- Lafayette	107
6. All Vehicle Trips per Household by Household Size and Auto Ownership -- Lafayette.	108
7. Home Based Work Vehicle Trips per Household by Household Size and Auto Ownership -- Evansville.	109
8. Home Based Other Vehicle Trips per Household by Household Size and Auto Ownership -- Evansville.	110
9. Non-Home Based Vehicle Trips per Household by Household Size and Auto Ownership -- Evansville.	111
10. All Vehicle Trips per Household by Household Size and Auto Ownership -- Evansville	112

HIGHLIGHT SUMMARY

Initiating a continuing, comprehensive and coordinated transportation planning process is a considerable task that faces small metropolitan areas as their population reaches 50,000. These areas usually lack the technical know-how as well as the financial capabilities needed to use the battery of computer-oriented planning models more commonly used in larger metropolitan areas. In most cases, such sophistication is not needed because the problems faced by smaller areas are different in nature and in complexity from those in larger areas.

This study examined the small and medium sized urban areas (population range: 50,000 to 250,000) of Indiana where full scale transportation studies have already been conducted, in an effort to make more efficient use of the information made available by these studies. This study focused on trip generation (frequency) models and parameters and examined the stability and transferability of these parameters between small and medium sized urban areas at three levels of aggregation: area-wide, zonal and household levels. A framework was thus provided for the transferability of trip frequency parameters. Special emphasis was placed on trip generation by category analysis.

Synthetic trip generation procedures aimed at eliminating the need for costly home interview O-D surveys were recommended.

A methodology based on Bayesian statistics was suggested for updating cross-classified household trip generation rates.

CHAPTER 1: INTRODUCTION

Initiating a continuing, comprehensive and coordinated transportation planning process can represent a formidable task for small metropolitan areas as their population reaches 50,000. This process usually follows the pattern set by the larger metropolitan areas in the 1950's (namely Detroit, San Juan and Chicago), with its high requirements in terms of data collection (Origin - Destination Survey), technical complexity and financial resources. However, the problems faced by the smaller areas are usually different in nature, in magnitude and in context from those in larger areas. Moreover, the changing emphasis in Urban Transportation Planning (1) from long-range large system planning to short-range improvements aimed at making better use of the existing facilities, coupled with the public concern over environmental and energy issues, have led to a reassessment of funding priorities, thus placing serious financial constraints on the development of "classical" full-scale comprehensive transportation studies. It is therefore imperative to develop and implement simplified alternative planning approaches which would greatly reduce the time and cost currently required by the "classical" transportation planning process, thus saving precious resources which can be redirected towards other short term and more pressing concerns.

One approach to simplifying the transportation planning process is to eliminate the need for a full scale Origin-Destination (OD)

survey which is by far the most costly, time consuming and time-delaying element of the conventional process. This can be achieved by reproducing the travel patterns in the urban area under study, based on the socio-economic and physical characteristics of that area, using parameters and relationships developed and calibrated in other areas where comprehensive OD surveys have been conducted. Such an approach is referred to as Synthetic Travel Demand Modeling, whereby the information traditionally obtained from the OD survey is fabricated or synthesized through the use of parameters "borrowed" from "similar" areas. However, questions arise as to which parameters and relationships can be transferred, as well as to which areas can be used as a source for such parameters.

This study focused on trip generation (frequency) models and parameters and examined the stability and transferability of these parameters between small and medium sized urban areas (population range: 50,000-250,000) at three levels of aggregation: areawide, zonal and household levels. The study was aimed at reducing the cost and complexity of the conventional planning procedures by developing a framework for transferring trip frequency parameters, thus eliminating the need for conducting full scale travel surveys.

In anticipation of such savings for future transportation studies, this research endeavour examined those small and medium urban areas of Indiana where full scale transportation studies have already been conducted, in an effort to make more efficient use of the information made available by these studies. In line with the above general aims, this study had the following primary objectives:

1. Compilation and interpretation of various travel characteristics for the several small metropolitan areas in Indiana in which origin-destination studies have been conducted.

2. Comparison of these travel characteristics and identification of the other urban variables that explain these characteristics.

3. Detailed examination of trip generation parameters and models for these areas at three levels of aggregation: areawide, zonal and household; assessment of the stability and transferability of these parameters for synthetic modeling applications.

4. Development of a generalized framework for the transferability of trip frequency (generation) parameters and models.

Some of the secondary objectives achieved by this study include:

1. Development of a data base which could be used for the following purposes:

- a. Cross-checking of the output of the planning process in a given area by comparing it to available information from other areas for the purpose of assessing its reasonableness (2).

- b. Input for quick-estimation techniques ("Quick Response" techniques) which might be needed for rapid evaluation of policy alternatives (3).

2. Determination of the number of levels needed within each variable in the cross-classification trip generation technique for small metropolitan areas.

3. Development of a statistical procedure for periodic updating of household trip generation rates.

Background and Scope of Study

The tools and techniques developed in the 1950's for large scale metropolitan transportation planning have become more or less standardized, and have been extensively used in a large number of urban areas. These procedures and models are thoroughly documented in the literature, and the necessary computer software is readily available from federal government sources. The modeling task is usually carried out in four stages: generation, distribution, mode split and traffic assignment. Each stage generates the input needed for the next stage. These models, in order to predict future travel, are usually calibrated so that they replicate base year travel patterns, with the implied assumption that the relationships thus established between travel and land use as well as socioeconomic characteristics of the population are stable over time. A vast amount of data is needed for this purpose; it is typically collected through an Origin-Destination survey consisting of a Home Interview Survey, Truck and Taxi Surveys as well as an External Cordon Survey. The Home Interview Survey is the largest and most costly of the three surveys, and its elimination is the target of most synthetic modeling approaches.

This large scale four-step modeling system has been subject to severe criticism in recent years. Some of this criticism has been directed at the deficiencies of the overall modeling system and the sheer size at which they operate (4,5,6,7), other criticisms have addressed problems with specific models (8,9,10), while others have questioned the basic behavioral assumptions underlying the whole modeling approach (11,12). Nevertheless, these models remain very widely

used, especially since professionals in the transportation planning area have become familiar with them.

The major issue however concerns the relevance of this large scale planning effort to the needs of small and medium metropolitan areas. These are more concerned with transportation costs and traffic congestion as opposed to the socio-economic and land use impacts which are of concern to the larger areas (13). Smaller areas therefore show a greater need for short-range Transportation System Management (TSM) type of planning, which would be of immediate benefit to the community, than for long range area wide type of planning, the results of which would be subject to a great deal of uncertainty. The appropriate level of effort for transportation planning activities depends on the size of an area and the complexity of its transportation problems. The Federal Highway Administration, Office of Highway Planning (14), has emphasized the need for "simplifying planning techniques, along with the scaling down of the reliance on sophisticated and data intensive computer techniques. Simplified approaches to planning will cut down on costs and time consumed in updating and reappraising long range plans, while allowing more time for focus on currently pressing demands such as TSM planning". It also suggested the elimination of large-scale O-D surveys and relying on small sample surveys if at all necessary.

Along the same line, a question is raised (15, p.18) as to "whether this increased level of detail, complexity and cost is justified by a corresponding increase in the accuracy and validity of the models relative to what could be achieved by a more simplified approach".

The need for the simplification of the transportation planning process in small urban areas is all too apparent. Various efforts have

been undertaken in this direction. Of special interest to this study are those aimed at eliminating O-D surveys by synthesizing travel patterns based on information from other areas, especially for the trip generation stage of the planning process. Grecco et al. (16) investigated the current organizational and technical aspects of transportation planning in small urban areas, and came up with recommendations about organizational improvements aimed at reducing the time and cost of the planning process. He also examined the use of several synthetic modeling approaches, such as the borrowing of zonal trip generation regression equations and of household trip rates developed in other small urban areas. These results will be discussed in greater detail when these particular techniques are examined in later chapters.

Coomer and Corradino (17) demonstrated the use of a short cut analysis technique for trip generation by travel purpose on an areawide basis in a small urban area. Total vehicle trips for the whole area were synthesized based on a relationship developed between total vehicle trips and area population based on observations from 14 other urban areas.

An extensive research effort in areawide travel forecasting was undertaken by Chan (18). He developed a classification scheme for all urban areas based on population and urban structure, and developed equations relating areawide trip frequency and areawide trip length to areawide auto ownership and area population respectively. Separate equations were developed for each urban area category.

The development and application of synthetic zonal trip generation equations has been undertaken by various researchers (19,20,21). One researcher, Bates (19), combined four sets of generation equations into

a single set of equations intended for use in the small urban areas of Georgia. This set was used in a limited number of areas for validation purposes, and the results, even though not totally convincing, were judged as satisfactory for the level of accuracy needed for these areas.

Recent efforts (22,23,24) have examined the use of cross-classification disaggregate models for generation analysis. This technique will be discussed in great detail in chapter six. Its use has been strongly suggested for synthetic modeling applications, especially in small urban areas (14). It consists of developing household trip rates for each household socio-economic category. Since the household is taken to be the basic unit of analysis, these models are behavioral in nature and should therefore be transferable from one area to another. Some evidence has been presented to that effect (24,25), while, on the other hand, some doubts have been expressed concerning the validity of this approach (16). The stability and transferability of household trip rates between small and medium urban areas is one of the matters examined in this study. No systematic and comprehensive treatment of the transferability of urban travel characteristics and especially trip frequency (generation) parameters at all levels of aggregation has been reported in the literature, and no criteria have been established to guide such a time and cost saving transfer.

In contrast to the previous work done in this topic area, this study has developed a framework for transferring travel frequency parameters at three levels of aggregation: areawide, zonal and household. This theoretical framework is supported by empirical evidence derived from the comparison of parameters obtained from various study areas for each of the relevant levels of aggregation.

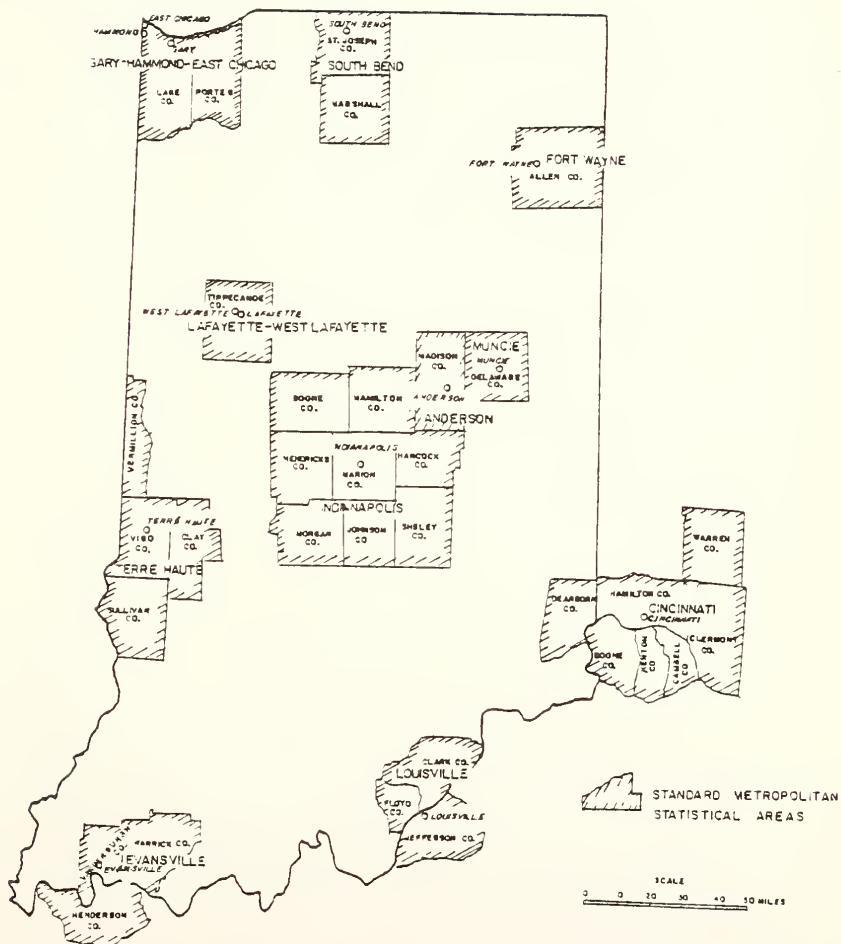
The study areas used for this analysis consisted of the small and medium sized metropolitan areas within the state of Indiana. The results of this study were intended for use in other urban areas within Indiana which are expected to reach a population of 50,000 in the near future. However, it is anticipated that the results and procedures are directly applicable to other midwestern communities that fall into the same size group. Moreover, because of the general nature of these results, their validity is by no means limited to the state or regional level.

CHAPTER 2: GENERAL CHARACTERISTICS OF THE STUDY AREAS

This study was aimed at compiling information from those metropolitan areas in the state of Indiana in which a detailed travel survey has been conducted and a transportation study performed. The focus of this study was small and medium sized areas with populations in the range of 50,000 and 250,000. Moreover, areas that are part of larger multi-state metropolitan areas were not included because of the inevitable special interaction within these regions which undoubtedly affects the transportation characteristics of these areas in a manner that might differentiate them from "typical" small or medium urban areas. Three such areas were encountered in Indiana: Gary-Hammond-East Chicago; Louisville; and Cincinnati. Indianapolis was the one other metropolitan area that was not included. This area was excluded because its population size exceeded the range examined in this study.

After the above eliminations, seven metropolitan areas remained. These are: Anderson, Evansville, Fort Wayne, Lafayette, Muncie, South Bend and Terre Haute. These Standard Metropolitan Statistical Areas (SMSA), as defined for U.S. Census purposes, are shown in Figure 1.

SMSA's are taken by the Bureau of Census to comprise one or more whole counties. However, parts of these counties might be largely underdeveloped and might thus not have been included in the study area for transportation planning purposes. For this reason, the boundaries of the SMSA's and those of the corresponding transportation study areas



STANDARD METROPOLITAN STATISTICAL AREAS - INDIANA

Source : U.S. Bureau of Census

FIGURE 1

do not necessarily coincide. The general characteristics of these areas, as reported in the transportation study reports, are shown in Table 1. Also reported is the year in which the O-D survey was conducted.

The information needed for this research effort consisted of the following:

- Transportation study reports or working papers generated by the planning process. Each area had its own different set of reports and working papers. One report in each set usually consists of a summary of the O-D data, and another report usually deals with travel model development, both being of vital importance to this study.
- Zonal level data: number of trips for each purpose produced by and attracted to each zone as well as the socio-economic and land use characteristics of each zone.
- Household level data, or the basic raw data obtained from the O-D survey. It usually consists of a set of trip records (purpose, time of day, mode,... of each trip) and a set of household socio-economic characteristics. This data was needed in order to relate the number of trips made for each purpose by each household sampled to the socio-economic characteristics of that household, and thus develop cross-classified trip rates. Due to the volume of this data, it is usually stored on magnetic tape for computer processing. One problem encountered, however, was that after the needed information had been extracted, the raw data files were simply deleted or purged by some of the study areas. This was especially true for those areas where the data was collected a

Table 1. Study Areas General Characteristics

	ANDERSON	EVANSVILLE	FORT WAYNE	LAFAYETTE	MUNCIE	SOUTH BEND	TERRE HAUTE
Survey Year	1971	1970	1961	1970	1971	1967	1971
Population	90,338	175,514	203,861	101,125	100,056	219,018	102,729
Occupied Dwelling Units	29,808	60,500	64,780	29,758	31,015	69,091	41,418
Auto per Household*	1.30	1.26	1.36	1.27	1.30	1.30	1.18
Auto per Household**	1.30	1.30	1.16	1.54	1.52	1.25	1.00
Persons per Household	3.03	2.90	3.15	3.40	3.33	3.17	2.48

* From Census Data for the SMSA -1970

** From Transportation Study Reports For Study Area (within Cordon Area)

long time ago. This is really unfortunate because that raw data is quite rich and would have been used for a variety of applications. However, such applications and analysis techniques were probably not anticipated at the time. As a consequence, such data was available for only two of the study areas: Evansville and Lafayette.

All of the above information was obtained either directly from the respective area planning commissions or through the Planning Division of the Indiana State Highway Commission.

CHAPTER 3: DISTRIBUTION OF TRIPS BY PURPOSE

Trips are usually stratified into three types in urban transportation planning studies. These are:

- Internal trips which are those with both ends within the study area cordon line.
- Internal-External trips, with one end inside the study area and the other outside the cordon.
- External-External or Through trips, with both ends outside the study area cordon.

The numbers and percentages of vehicle trips of each type made on a "typical" weekday are shown in Table 2. The percentages of each type vary from one area to another, depending on the regional role of each of the cities, and on its location with respect to intercity travel corridors (for through trips). However, another important factor in the percentage of trips of each type is the location of the cordon line which delineates the boundaries of the study area. A certain number of trips could be classified as either internal or external depending on where the cordon line is drawn.

On the average, 82.2 percent of the vehicle trips are internal, 15.2 percent are external-internal and 2.6 percent are through trips. These percentages are the (unweighted) averages of the respective percentages for each of the seven urban areas, as reproduced in Table 2.

Table 2. Number and Percentage of Vehicle Trips
by Type and Urban Area

<u>Urban Area</u>	<u>Internal</u>	<u>Internal- External</u>	<u>Through</u>	<u>Total</u>
ANDERSON	290,000 (78.0)	64,000 (17.2)	18,000 (4.8)	372,000 (100.0)
EVANSVILLE	452,794 (87.4)	60,206 (11.6)	5,279 (1.0)	518,279 (100.0)
FORT WAYNE	387,222 (83.7)	67,240 (14.5)	8,341 (1.8)	462,803 (100.0)
LAFAYETTE	295,710 (84.4)	44,034 (12.6)	10,638 (3.0)	350,382 (100.0)
MUNCIE	288,249 (79.8)	69,530 (19.2)	3,674 (1.0)	361,453 (100.0)
SOUTH BEND	530,020 (84.3)	86,177 (13.7)	12,300 (2.0)	628,497 (100.0)
TERRE HAUTE	233,711 (77.7)	53,022 (17.6)	14,240 (4.7)	300,973 (100.0)
Unweighted Average Percentages	(82.2)	(15.2)	(2.6)	(100.0)

(Numbers in parentheses are percentages.)

Internal trips are by far the most important (number-wise) among the three types, and their forecasting constitutes the bulk of the transportation planning process.

Internal trips for the base year have been traditionally obtained by the home-interview survey, while external trips (both external-internal and through trips) have been determined from the external cordon survey. As explained in the first chapter, the home interview survey is the most costly and time consuming element of the Origin-Destination survey. Moreover, in as much as the aim of this research effort is to examine synthetic techniques which would minimize or eliminate the need for the home interview survey, it essentially focuses on internal trips; information on external trips would be available from the external cordon survey which would still be conducted for a synthetic urban transportation study. Internal trips will therefore be discussed in greater detail, and the variation of their distribution by purpose between urban areas as well as within urban areas will be examined in the following sections.

Number and Definition of Trip Purposes

Transportation studies have usually stratified internal trips into as many as seven purposes. In the interest of simplification, it is recommended that fewer trip purposes be used in the demand modeling process (14). For smaller urban areas, the following three trip purposes are usually adequate: Home-Based Work (HBW), Home-Based Other (HBO) and Non-Home Based (NHB).

Home-Based Work trips are those made between a person's residence and his place of employment. Trips originating at home and ending at

work as well as those originating at work and ending at home are considered as HBW trips.

Home-Based Other trips are those made between a person's residence and any place different from his place of employment. These trips are often further classified into these purposes: School, Personal Business, Shopping, Social-Recreation, etc. For smaller areas however, aggregation into one category (HBO) has proven to be quite acceptable. As in the case of HBW trips, either origin or destination should be the traveller's home.

Non-Home Based trips are those which have neither trip end (neither origin nor destination) at the traveller's residence. For smaller urban areas where truck and taxi trips constitute a small percentage of the internal trips, these trips are included in the NHB category; this is a standard procedure used in most small area transportation studies.

These three purposes as defined above have been adopted for this research and are used throughout this study.

Variation Between Areas

The distribution of internal vehicle trips by purpose for each of the study areas is shown in Table 3. For those areas where more than three purposes were used in the transportation study, the trips were combined into the three categories described above.

Wilson and Kristoffersen (20) have shown that trip purpose distribution is independent of city size for smaller and medium sized cities. In their study, three separate regression equations were developed relating the percentage of HBW, HBO and NHB trips respectively to city population. As an example, the equation for percentage HBW

Table 3. Distribution of Vehicle Trips by Purpose

<u>City</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB*</u>	<u>TOTAL</u>
ANDERSON	47,105 (16.2)	145,958 (50.3)	97,139 (33.5)	290,202 (100)
EVANSVILLE	74,586 (16.5)	188,285 (41.6)	189,923 (41.9)	452,794 (100)
FORT WAYNE	90,203 (23.3)	146,990 (38.0)	150,029 (38.7)	387,222 (100)
LAFAYETTE	44,337 (15.0)	154,067 (52.1)	97,306 (32.9)	295,710 (100)
MUNCIE	38,591 (13.4)	159,017 (55.2)	90,641 (31.4)	288,249 (100)
SOUTH BEND	79,672 (15.1)	208,452 (39.3)	241,896 (45.6)	530,020 (100)
TERRE HAUTE	36,745 (15.7)	121,836 (52.1)	75,130 (32.2)	233,711 (100)

*Including Truck trips

(Numbers in parentheses are percentages)

Unweighted Average Percentages: HBW: 16.5%

HBO: 46.9%

NHB: 36.6%

trips was: $HBW (\%) = 22.5 - 0.00001 \times \text{Population}$. The coefficient of the independent variable (population) in each equation was tested for significance (using a 95% confidence interval based on t-distribution). These tests showed that this coefficient was not significant for any of the purposes.

The validity of the inferences made above should be seriously questioned. First of all, the regression model developed included only one independent variable, population, without statistical or theoretical proof that it was uncorrelated with other potential independent variables that could have significantly affected the dependent variable (percentage of trips for each purpose). Thus the regression coefficient of population would not reflect any inherent effect of this particular independent variable on the dependent variable because its magnitude is affected by the correlated independent variables not included in the model (26, pp.252,346). Therefore, the conclusion that trip purpose distribution is independent of size might not be justified. Moreover, even it were, it would not necessarily imply that this distribution is "identical" for all cities.

The hypothesis of the independence of the distribution of trips among purposes from the various factors that differentiate cities is tested in this section. The Chi-Square test is considered to be an appropriate technique for this purpose. The appendix describes the details of this testing technique which will be used in various other instances in this study. One of the reasons for using this technique is that it does not place any requirements on the distribution from which the observed variables are sampled. In addition, the Chi-Square

test is suitable for use with categorical data such as that being analyzed.

In order to use this test, the number of trips by purpose and by urban area can be arranged in a contingency table, with "Urban Area" as the row factor and "Purpose" as the column factor. Since there are seven urban areas and three trip purposes, a 7 x 3 contingency table is obtained, with 21 possible and mutually exclusive combinations or cells. To each one of these cells corresponds an observed frequency (from the sample), which in this case is the number of trips directly sampled in that specific area for that specific purpose. This last point is very important because it means that the actual number of trips obtained from the O-D survey (before expansion) should be used instead of the numbers shown in Table 3 (which were expanded from the O-D surveys) or of the percentages also shown in that table. This is due to the fact that the test statistic is sensitive to sample size since the test tries to assess whether the differences observed for each of the urban areas could have occurred merely as the result of sampling error. For this analysis, unexpanded survey data was not available for all of the study areas, neither was the sampling rate used for the survey. For full-scale O-D surveys, sampling rates are usually found to be in the range of 8-12.5 percent. A flat 10 percent sampling rate was assumed for all of the study areas in order to develop the contingency table for the Chi-Square test. Even though this assumption is not entirely correct because the sampling rate was different in each of the areas, it is quite adequate for this test, as evidenced by the magnitude of the test statistic ($\chi^2_{\text{data}} = 5596$) vs. $\chi^2_{12, 0.005} = 28.299$) which is so large

that any "refinements" in the values of the sampling rate will not change the outcome of the test.

The contingency table as well as the test statistic calculations are summarized in Table 4. The hypothesis of independence of trip purpose distribution from "urban area" is rejected with 99.5% confidence. The results of this test therefore demonstrate that the distribution by purpose of internal vehicle trips is not the same for all the small urban areas in Indiana.

Pairwise comparisons of each of the urban areas using Chi-Square tests were also conducted. 2 x 3 contingency tables (2 urban areas and 3 purposes) were set up for each possible pair of study areas. Chi-Square tests provided statistical evidence for the rejection of the hypothesis that the trip distribution by purpose is identical in any two of the urban areas studied. The only pair of urban areas for which this hypothesis could have been accepted consisted of Lafayette and Terre Haute. That particular contingency table as well as the value of the test statistic are shown in Table 5. This Table reveals that the hypothesis could be rejected with 95% confidence but not with 97.5% confidence. In order to generalize the results of these comparisons, the hypothesis is rejected for Lafayette and Terre Haute. For all the other comparisons, the hypothesis is rejected with 99.5% confidence.

It should be noted that the Chi-Square test is very sensitive to small differences in the observed frequencies, especially for large samples such as the one examined here. This can be seen in the Lafayette and Terre Haute comparison. These two areas have only very slightly different percentages of trips for each purpose (cf. Table 3) yet the

Table 4. Summary of Chi-Square Test -
Effect of Urban Area on Trip Purpose Mix

<u>Urban Area \ Purpose</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB</u>
ANDERSON	4,710	14,596	9,714
EVANSVILLE	7,459	18,828	18,992
FORT WAYNE	9,020	14,699	15,003
LAFAYETTE	4,434	15,407	9,731
MUNCIE	3,859	15,902	9,064
SOUTH BEND	7,967	20,845	24,190
TERRE HAUTE	3,674	12,184	7,513

H_0 = Distribution of trips by purpose is independent of Urban Area.

$\alpha = 0.005$

Test Statistic: $\chi^2_{data} = 5596$

$$\chi^2_{12,0.005} = 28.299 < \chi^2_{data}$$

Conclusion: Reject H_0

Table 5. Chi-Square Test of Similarity of Distribution of
Trips by Purpose Between Lafayette and Terre Haute.

<u>Urban Area \ Purpose</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB</u>
LAFAYETTE	4,434	15,407	9,731
TERRE HAUTE	3,674	12,184	7,513

H_0 : Distribution of trips by purpose is identical in Lafayette and Terre Haute

$$\chi^2_{data} = 6.81$$

$$\nu = 2$$

For $\alpha = 0.05$, $\chi^2_{2,0.05} = 5.991 < \chi^2_{data}$: Reject H_0

For $\alpha = 0.025$, $\chi^2_{2,0.025} = 7.378 > \chi^2_{data}$: Accept H_0

hypothesis of identical distributions was rejected with 95% confidence. Therefore, the results of these tests should not discourage synthetic modeling efforts consisting of borrowing relationships and parameters from "similar" areas. For one, no two urban areas are identical, and the statistical evidence presented above was expected. Every urban area has its own unique set of characteristics: form, structure, socio-economic mix of its residents, economic base, activity systems, etc. Moreover, synthetic modeling techniques for small urban areas do not assume statistically identical cities, and are only intended as alternative planning tools offering sufficient reliability at a significantly lower cost than those requiring full-scale O-D surveys. However, these tests do indicate that careful judgement should be used when borrowing and transferring travel parameters and relationships. Thus, it is of primary importance to understand what parameters can be transferred, at what level (of aggregation) and under what circumstances; these are the major questions addressed by this study as far as the frequency or generation aspect of tripmaking is concerned.

Use of Statewide Averages

A major factor in the choice of analysis tool to aid in decision-making is the issue addressed and the sensitivity of the anticipated decision to the information generated by the analysis. This in turn determines the degree of reliability and accuracy needed in the analysis output. Considering the relatively large uncertainty accepted in transportation studies due to the very nature of the issues addressed (especially in the context of smaller urban areas where the decision concerns the number of lanes of a certain facility), the trip purpose

distribution percentages obtained for the study areas shown in Table 3 could be quite useful.

A "typical" trip purpose distribution can be obtained by averaging the percentages of each purpose for all the study areas. Since only approximate descriptive values are desired, simple unweighted averages are computed for each purpose. These are:

HBW : 16.5%

HBO : 46.9%

NHB : 36.6%

These average percentages could be used for developing gross estimates, or "macro-estimates" (2) which can be used as reasonableness cross checks on areawide totals, or as approximations in rough first cut evaluations by decision makers and administrators. This average distribution could also be used as an initial assumption (to be adjusted later in the process) in a synthetic modeling effort (27).

Distribution of Person Trips by Purpose

All of the above numbers and percentages are for vehicle trips. Most of the transportation studies considered only vehicle trips due to the very low and almost insignificant percentage of transit trips in the small urban areas of Indiana. No testing will therefore be made for person trips, especially since appropriate data is not available for some of the study areas. However, figures for person trips will be shown when available in order to convey a clearer picture of the travel patterns in the study areas. The distribution of person trips by purpose for some of the study areas is shown in Table 6. The comparison of this Table with Table 3 (vehicle trips) shows that the percentage of

Table 6. Distribution of Person Trips by Purpose

<u>Urban Area</u>	<u>Purpose</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB</u>	<u>TOTAL</u>
ANDERSON		56,752 (14.4)	249,037 (63.3)	87,406 (22.3)	393,195 (100.0)
EVANSVILLE		94,965 (15.7)	338,940 (56.2)	169,697 (28.1)	603,602 (100.0)
MUNCIE		45,674 (11.4)	258,617 (64.4)	97,430 (24.2)	401,721 (100.0)
SOUTH BEND		102,853 (15.2)	354,235 (52.5)	218,168 (32.3)	675,256 (100.0)

Unweighted Average Percentages: HBW : 14.2%

HBO : 59.1%

NHB : 26.7%

(Numbers in brackets are percentages)

HBO person trips is consistently higher than that of HBO vehicle trips, while the percentages of total person trips for the other two purposes are lower than the corresponding percentages of vehicle trips. This is a reflection of the fact that auto occupancies for HBO trips are higher than those for the other two purposes (28). Overall, the same variation of the trip purpose distribution between the study areas for vehicle trips is observed for person trips.

Variation Within Urban Areas

In order to understand some of the factors behind the variation of the trip purpose distribution between urban areas, this distribution was investigated within urban areas. A disaggregate level of analysis was adopted for this purpose, and trips made by individual households were considered. Separate trip purpose distribution models were developed for each socio-economic household category, and a statistical

test of the independence of the distribution of trips by purpose for households and the socio-economic characteristics of these households was conducted.

In order to perform the above test, disaggregate household data was needed. O-D surveys would commonly provide such data. For this study, a good set of raw O-D data collected in Evansville was available, and was therefore used for this analysis. This data was processed to yield the number of vehicle trips for each purpose made by the members of each household surveyed. Valid information was available for 5507 such households.

The next step was to group trips made by households on the basis of the socio-economic characteristics of the households. Auto ownership was used as an indicator of these characteristics as suggested in the literature (22), especially since it is highly correlated with trips made by households as well as with household income (8). Moreover, its ease of forecasting makes it a very desirable socio-economic variable for travel demand analysis. Auto ownership information was readily available from the O-D survey for each of the households sampled.

Households were therefore stratified by auto ownership category. Four categories were formed. The number of households in each of these categories was as follows:

<u>Auto Ownership</u>	<u>Number of Households</u>
0	862
1	2230
2	2018
3 or more	<u>397</u>
	Total: 5507

Survey trips for each purpose were accumulated within each auto ownership group. This was done using the BREAKDOWN procedure of the SPSS (Statistical Package for the Social Sciences) package available at Purdue University's computing facilities. The number of trips thus obtained for each purpose and car ownership level (category) was then used to compute the percentage distribution of trips by purpose for households within each auto ownership group. These percentages are shown in Table 7. This distribution is also plotted in Figure 2.

In order to test the hypothesis that the distribution by purpose of trips made by households was independent of the socio-economic characteristics of households (as reflected in the auto ownership status), a Chi-Square test similar to the one described earlier was used. General computational details are shown in the Appendix. A 4 x 3 contingency table was developed with "Auto Ownership" as the row factor and "Purpose" as the column factor. The frequencies observed for each combination of factor levels consisted of the number of surveyed trips for each purpose within each household car ownership group. The contingency table as well as the Chi-Square test for independence are summarized in Table 8. As shown in that table, test results led to the rejection of the independence hypothesis with 99.5% confidence. Therefore, the distribution of vehicle trips by purpose turned out to be significantly different between auto ownership groups. Auto ownership in this case was used only as a descriptive measure of socio-economic status and not as an exclusive factor explaining the split of trips by purpose. For predictive purposes, finer stratifications would lead to more accurate results. However, the main concern at this stage is to show the variation within urban areas due to the socio-economic characteristics of

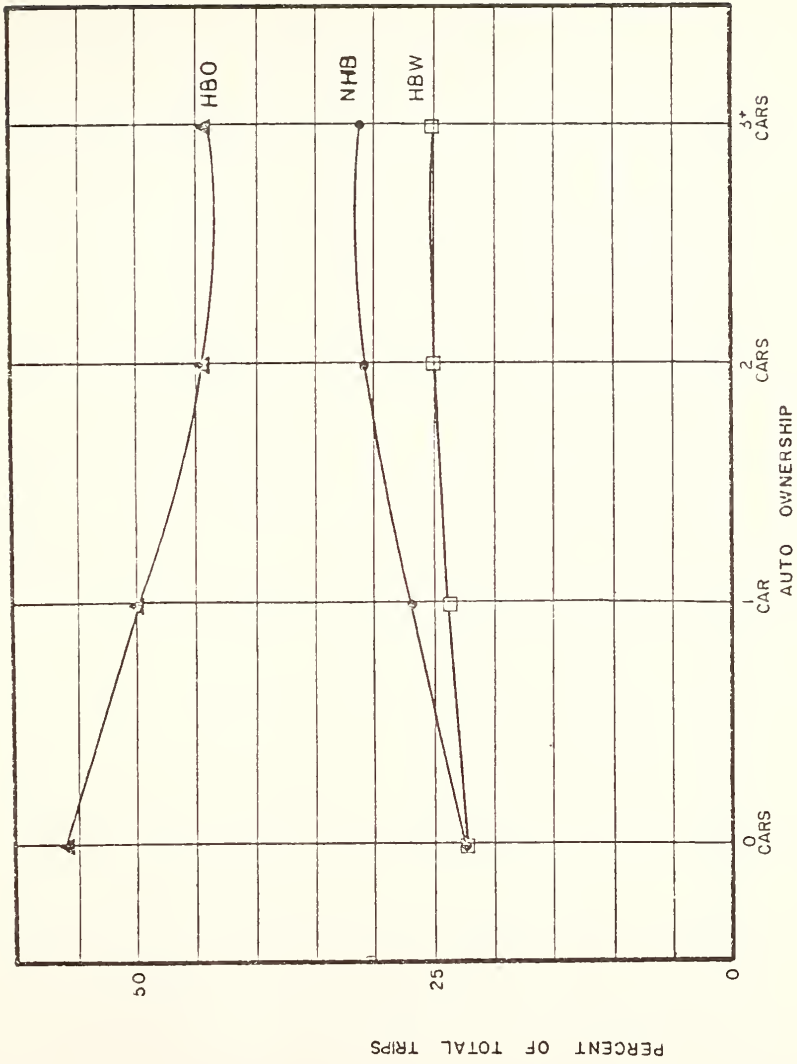


FIGURE 2 : PURPOSE DISTRIBUTION OF VEHICLE TRIPS BY AUTO OWNERSHIP - EVANSVILLE

Table 7. Distribution of Household Vehicle Trips by
Purpose and Household Auto Ownership--EVANSVILLE

<u>Purpose</u> <u>Auto Ownership</u>	<u>%</u> <u>HBW</u>	<u>%</u> <u>HBO</u>	<u>%</u> <u>NHB</u>
0	22.2	55.6	22.2
1	23.5	49.9	26.6
2	25.2	44.1	30.7
3+	25.0	44.1	30.9

Table 8. Contingency Table and Summary of Chi-Square Test of
Independence of Trip Purpose Distribution and Auto Ownership--EVANSVILLE

<u>Purpose</u> <u>Auto Ownership</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB</u>	<u>TOTAL</u>
0	6	15	6	27
1	2,351	4,996	2,661	10,008
2	4,000	6,997	4,872	15,869
3+	1,117	1,974	1,381	4,472

H_0 : distribution by purpose of trips made by households is
independent of auto ownership level of households

$$\chi^2_{\text{data}} = 97$$

$$\text{d.f.} = 3 \times 2 = 6$$

$$\chi^2_{6,0.005} = 18.548 < \chi^2_{\text{data}}$$

Conclusion: Reject H_0

the households, and to relate this variation to the "between urban areas" variation as explained in the next section.

In as much as the trip distribution by purpose is significantly different for each socio-economic group, it would appear reasonable to assume that the overall (areawide) trip distribution by purpose in two different urban areas would be the same, if the socio-economic mix (distribution of households among socio-economic groups) in these areas were similar. This implies that "borrowing" areawide percentages without first comparing the socio-economic mix of the two areas (source and "borrower") could lead to erroneous results. For synthetic applications, borrowing and transferring separate relationships for each socio-economic group is always preferred. Even then, a great deal of care ought to be exercised since a great deal of variability is bound to exist within each socio-economic group, in addition to the fact that the socio-economic indicator(s) used might not capture all of the causative effects.

The appropriate level of aggregation at which parameters and models could be transferred will be thoroughly investigated and described in greater detail in later chapters where a comparative refined analysis of trip frequency and its forecasting techniques will be discussed.

In the specific case of trip purpose distribution models stratified by socio-economic categories, it was suggested that these could be transferred to areas of "similar character and size"(22). For small urban areas transportation planning, synthetic applications using the models developed above could probably provide sufficient reliability. However, it is felt that this synthetic technique is not the

most appropriate one for detailed studies, considering the fact that only one socio-economic variable (auto-ownership) has been used. Therefore, no strong case is made for the "universal" transferability of these models. However, they would give better reliability than areawide aggregate purpose distributions. Later chapters will present alternative approaches which are more appropriate for synthetic detailed forecasting activities.

CHAPTER 4: AREAWIDE TRIP FREQUENCY

This chapter deals with trip frequency or generation at the area-wide level of analysis. It shows various descriptive indices for each of the Indiana study areas and compares them to nationwide data. It also identifies some of the factors underlying areawide travel frequency and presents a comparative examination of some of these factors for each of the study areas. Implications for transferability and synthetic applications are also discussed.

In order to place matters into the proper perspective, a definition of the term "areawide" as used in this context is necessary. "Areawide" describes the geographic unit of analysis which in this case is the totality of the urban area for which a transportation study is being conducted. It implies that the representative measures adequately reflect the whole urban area, and that the results anticipated are figures for the whole area. Moreover, these areawide frequency measures are to be deduced based on aggregate areawide characteristics of the urban area. This is the highest level of aggregation considered in this study. This level is therefore "areawide" in two respects:

- Trip frequency responses are considered for the whole urban area.
- These measures are "predicted" or assumed based on areawide characteristics of the urban area. In other words, models for the prediction of areawide trip frequency responses could be calibrated using areawide independent variables.

The following rates are the most commonly used descriptive measures of trip frequency:

- Average number of trips per household (per day), which can be obtained by dividing the total number of trips by the total number of households in the study area.
- Average number of trips per person (per day), which is obtained by dividing the aggregated number of trips for the area by the total population.

The above two rates reflect the trip making activity of a "typical" household and of a "typical" person respectively in the study area. Another rate which is sometimes used is the "average number of trips per automobile".

Table 9 shows the average vehicle trip rate per household for each purpose for each of the study areas. This rate is for internal vehicle trips only. Also shown in Table 9 are "typical" rates obtained from national data (28, p.34). These rates exhibit a fairly wide range of variation between the study areas, with the largest range occurring for HBO trips (2.27 trips per household in Fort Wayne to 5.18 trips per household in Lafayette). The rate for all trip purposes is lowest in Terre Haute (5.64) and highest in Lafayette (9.94). The "mean" rates for all seven study areas are also shown in Table 9.

The average internal vehicle trip rate per person (or per capita) for each purpose is shown in Table 10 for each of the study areas. Mean rates are also shown. The variation of per capita rates between the study areas is not identical to that of the household trip rates considering that the average number of persons per household (reported in Table 1) is not the same for all the areas. Per capita rates exhibit

Table 9. Areawide Internal Vehicle Trip Rates:

Daily Trips Per Household

<u>URBAN AREA</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB*</u>	<u>ALL</u>
ANDERSON	1.58	4.90	3.26	9.74
EVANSVILLE	1.23	3.11	3.14	7.48
FORT WAYNE	1.39	2.27	2.32	5.98
LAFAYETTE	1.49	5.18	3.27	9.94
MUNCIE	1.17	4.83	2.92	8.92
SOUTH BEND	1.15	3.02	3.50	7.67
TERRE HAUTE	0.89	2.94	1.81	5.64
(Unweighted) AVERAGE	1.27	3.75	2.89	7.91

Typical Rates**

<u>Urban Area Population</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB</u>	<u>ALL</u>
50,000-100,000	1.68	4.44	2.06***	8.18
100,000-250,000	2.24	4.63	2.25***	9.12

* Includes truck trips

** Adapted from Reference 28, p.34. These "typical" rates are for the sum of Internal and External trips made by area residents.

***Does not include truck trips

Table 10. Areawide Internal Vehicle Trip Rates:

Daily Trips Per Capita

<u>URBAN AREA</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB*</u>	<u>ALL</u>
ANDERSON	0.52	1.62	1.08	3.22
EVANSVILLE	0.42	1.07	1.08	2.57
FORT WAYNE	0.44	0.72	0.74	1.90
LAFAYETTE	0.44	1.52	0.96	2.92
MUNCIE	0.39	1.59	0.91	2.89
SOUTH BEND	0.36	0.95	1.11	2.42
TERRE HAUTE	0.36	1.19	0.73	2.28
(Unweighted)Average	0.42	1.24	0.94	2.60

*includes truck trips

a relatively narrower range of variation between the study areas. Fort Wayne has the lowest trip rate (all purposes) while Anderson has the highest (1.90 and 3.22, respectively). It has to be kept in mind, however, that the data for Fort Wayne was collected in 1961 which is much earlier than the surveys conducted in the other study areas.

Table 11 shows the average internal vehicle trip rate for each purpose per automobile owned by the residents of each of the study areas, as well as mean rates for the seven study areas. This rate exhibits somewhat less variability between study areas than the two described above. The highest all-purposes rate per vehicle is in Anderson (7.55) while the lowest is in Fort Wayne (5.16).

The average household trip rates are presented in Table 12 for internal person trips as opposed to vehicle trips for some of the study areas. The "typical" rates (28, p.34) also shown are for internal trips as well as external trips and are therefore not directly comparable to

Table 11. Areawide Internal Vehicle Trip Rates:

Daily Trips Per Automobile

<u>URBAN AREA</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB*</u>	<u>ALL</u>
ANDERSON	1.22	3.80	2.53	7.55
EVANSVILLE	0.95	2.39	2.41	5.75
FORT WAYNE	1.20	1.96	2.0	5.16
LAFAYETTE	0.97	3.36	2.12	6.45
MUNCIE	0.97	4.00	2.28	7.25
SOUTH BEND	0.92	2.41	2.80	6.13
TERRE HAUTE	0.89	2.94	1.82	5.65
(Unweighted)AVERAGE	1.02	2.98	2.28	6.28

*includes truck trips

Table 12. Areawide Internal Person Trip Rates:

Daily Trips Per Household

<u>URBAN AREA</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB*</u>	<u>ALL</u>
ANDERSON	1.90	8.35	2.93	13.18
EVANSVILLE	1.57	5.60	2.81	9.98
MUNCIE	1.47	8.34	3.14	12.95
SOUTH BEND	1.49	5.13	3.16	9.78
(Unweighted)AVERAGE	1.61	6.86	3.01	11.47

Typical Rates**

<u>URBAN AREA POPULATION</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB</u>	<u>ALL</u>
50,000 - 100,000	2.24	8.54	3.22	14.00
100,000 - 250,000	2.80	7.98	3.22	14.00

* Does not include truck trips.

**Adapted from Reference 28, p.34. These "typical" rates are for the sum of Internal and External trips made by area residents.

the others. However, the order of magnitude is the same. NHB rates exhibit the least variability between the study areas and compare very well with the "typical rates". Per capita person trip rates, shown in Table 13, follow the same trend as household trip rates. NHB trip rates are strikingly stable over the different study areas. Most of the overall variability is due to the variability in HBO trips which average 2.22 trips per person for the study areas. The number of total trips per person for the study areas, 3.73, is higher than the national figures. A "reasonable" range for this rate has been reported to be 2.5 to 3.0 trips per person per day, of which approximately 0.6 to 0.7 trips per person per day are for work (2). HBW trip rates for most of the study areas are somewhat less than the lower limit of this range, with the exception of Anderson's (0.63). These ranges are, however, based on data collected in larger cities.

Table 14 contains the areawide number of internal person trips per automobile. The same trends observed earlier are still present for this parameter (stable NHB rates and fluctuating HBO rates). Anderson still has the highest overall rate (10.15), which is quite higher than the national "typical" rate of about 7 trips per car. Again the reader is reminded that these "typical" rates are derived from figures based on data from essentially larger cities (2).

The above detailed comparison of areawide trip rates leads to the conclusion that these rates do vary between urban areas, despite the similarity exhibited by some of them (especially Anderson and Muncie). This variability is to be expected, considering that areawide travel demand reflects various urban area characteristics which are most likely different between the study areas.

Table 13. Areawide Internal Person Trip Rates:

Daily Trips Per Capita

<u>URBAN AREA</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB*</u>	<u>ALL</u>
ANDERSON	0.63	2.76	0.97	4.36
EVANSVILLE	0.54	1.93	0.97	3.44
MUNCIE	0.46	2.58	0.97	4.01
SOUTH BEND	0.47	1.62	1.00	3.09
(Unweighted)AVERAGE	0.53	2.22	0.98	3.73

*Does not include truck trips

Table 14. Areawide Internal Person Trip Rates:

Daily Trips Per Automobile

<u>URBAN AREA</u>	<u>HBW</u>	<u>HBO</u>	<u>NHB*</u>	<u>ALL</u>
ANDERSON	1.46	6.43	2.26	10.15
EVANSVILLE	1.21	4.31	2.16	7.68
MUNCIE	1.15	6.5	2.45	10.10
SOUTH BEND	1.19	4.10	2.53	7.82
(Unweighted)AVERAGE	1.25	5.34	2.35	8.94

*Does not include truck trips

Factors Affecting Areawide Travel Frequency

The major determinants of urban travel demand are:

1) Urban Form, which has many dimensions, namely (29):

- Shape of the urban area, which was examined in several studies essentially concerned with the transportation requirements for each of the shapes (30,31).

- Spatial distribution of the various land uses in the urban area, in terms of their location and intensity. Studies have been undertaken in order to identify the travel requirements associated with various existing and hypothetical schemes for clustering residences, work and shopping locations and the various land uses (31,32,33).

- Urban area size, which was found to significantly affect urban auto trip demand (as measured by household trip frequency), with this demand being higher in smaller urban areas than in larger ones (29).

- Transportation network layout and characteristics.

2) Social and economic characteristics of the residents. Extensive empirical evidence has been presented to demonstrate the dependence of household tripmaking (and especially trip frequency) and the socio-economic characteristics of the household, especially auto ownership and household size (34). In as much as areawide travel is an aggregation of individual household travel, those characteristics affect areawide trip frequency rates. However, it is not necessarily true that these rates are a function of aggregate socio-economic variables (such as average auto ownership and average household size) for the whole area, since

these might not be truly representative of the distribution of these variables among the individual households. Chan (18), in an effort to relate areawide trip rates to areawide socio-economic variables, indeed found that while "average auto ownership per dwelling unit" (over the whole study area) showed the highest significant correlation with the trip rates, the "average number of persons per household" did not show significant correlation with the trip rates on an areawide basis (even though it is highly significant at the household level).

Synthetic Forecasting of Areawide Travel Frequency

The variation of areawide travel demand between urban areas therefore reflects the difference in the elements described above between these areas. Hence, synthetic derivation of areawide aggregate trip rates has to take these factors into account. A procedure for forecasting the average number of person trips per household (per day) for urban areas was developed by Chan (18). It consists of separate regression equations for each category of urban areas. These categories define a classification scheme based on the following factors:

- size of the urban area (in terms of population), for which two levels were defined: less than 800,000 and greater than 800,000.
- urban structure, which describes the spatial distribution of activities. Two states were used for this factor: core-concentrated and multi-nucleated.

This classification scheme should account for differences in urban form, including the characteristics of the transportation supply network. The regression equation for each category predicts household person trip frequency rate as a function of average household auto

ownership. The proposed classification scheme as well as the equations are shown in Table 15. Along with each equation is also shown the number of observations used to calibrate the equation as well as the coefficient of determination R^2 (measure of the goodness of fit of the regression equation; ranges between 0 and 1; the closer to 1 the better).

Three categories were eventually determined: Large and Core-Concentrated, Large and Multi-Nucleated, and Small (and Medium). This last category reflects the fact that for urban areas with population between 50,000 and 800,000, urban structure is not a differentiating factor when predicting areawide trip frequency rates. This result provides yet another argument in favor of simplified synthetic techniques for transportation planning in small and medium urban areas, since urban form does not seem to be a significantly differentiating factor between these areas as far as trip frequency is concerned. This is especially true for the study areas of Indiana (and the Midwest in general) where no major urban form differences exist between these areas.

The equation developed by Chan for this group of urban areas is:
 Average Number of Person Trips/Household = $1.262 + 6.591 \times (\text{Average Auto Ownership/Household})$

This equation is based on 81 observations and has a coefficient of determination (R^2) of 0.412.

Table 16 shows a comparison between the household person trip frequency rates observed in several of the study areas and those predicted using the above equation. Overall, the predicted values provided a good approximation of the observed values. However, one inherent weakness is apparent in these few observations: Anderson and

Table 15. Relationship Between Average Household Person Trip Frequency and Average Auto Ownership Per Household Stratified by Urban Size and Urban Structure

Population Size	Urban Structure	
	Core-Concentrated	Multi-Nucleated
Large with Population over 800,000	$\bar{Y} = 2.242 + 3.761 \bar{X}$ $R^2 = 0.824$ $n = 8$	$\bar{Y} = 2.831 + 4.117 \bar{X}$ $R^2 = 0.723$ $n = 9$
Small and Medium with Population between 50,000 and 800,000	$\bar{Y} = 1.262 + 6.591 \bar{X}$ $R^2 = 0.412$ $n = 81$	

where:

$\bar{Y}$ = average number of daily trips per dwelling unit

$\bar{X}$ = average auto ownership per dwelling unit

n = number of observations

R^2 = coefficient of determination

Source: Reference 18, p. 42.

Table 16. Comparison of Areawide Household Person Trip
Frequency Rates: Predicted vs. Observed

Equation used for prediction (Ref. 18)

Avg. Household Person Trip Frequency Rate =

$$1.262 + 6.591 \times (\text{Avg. Auto Ownership/Household})$$

<u>Urban Area</u>	<u>Avg. Auto per Household</u>	<u>Predicted Rate</u>	<u>Observed Rate</u>
ANDERSON	1.30	9.83	13.18
EVANSVILLE	1.30	9.83	9.98
MUNCIE	1.52	11.28	12.95
SOUTH BEND	1.25	9.5	9.78

Evansville have the same average auto ownership per household yet their average household person trip frequency rates are quite different. Therefore, as anticipated from the low R^2 value, extremely accurate predictions should not be expected from this equation. Chan's equation does provide a reasonable approximation of the trip rates. However, a comparable trip rate could also be obtained from using results from other areas or general "guidelines". Moreover, any areawide aggregate regression model is likely to suffer averaging biases, in as much as the "average auto ownership per household" might not be a good representative of the distribution of the auto ownerships for each individual household in the urban area. This distribution is investigated in the next section.

Comparative Analysis of the Household Auto Ownership Distribution

The information needed for this analysis was not available from the transportation studies for all the study areas. However, the U.S. Census reports the number of households in each auto ownership category for all Standard Metropolitan Statistical Areas (SMSA's). Even though the boundaries of the SMSA's and those delimiting the area included in the transportation study (cordon area) do not necessarily coincide (as explained in Chapter 2), it is felt that using census data for the whole SMSA serves the needs of this particular analysis.

The distribution of households among auto ownership categories in the SMSA corresponding to each of the study areas as obtained from 1970 Census data is shown in Table 17. Four auto ownership levels (categories) are distinguished: 0 cars, 1 car, 2 cars, 3 or more cars.

Table 17. Household Distribution by Auto Ownership

Number of Automobiles Available	ANDERSON	EVANSVILLE	FORT WAYNE	LAFAYETTE	MUNCIE	SOUTH BEND	TERRE HAUTE
0	4,852 (10.98)	11,701 (15.63)	9,480 (11.06)	3,602 (11.16)	4,521 (11.47)	10,617 (12.25)	9,612 (16.60)
1	23,609 (53.43)	35,842 (47.88)	40,938 (47.76)	17,990 (55.75)	20,382 (51.69)	44,632 (51.49)	30,699 (53.01)
2	13,663 (30.92)	23,548 (31.46)	29,980 (34.97)	9,043 (28.03)	12,485 (31.66)	26,565 (30.65)	15,029 (25.95)
3+	2,066 (4.67)	3,764 (5.03)	5,322 (6.21)	1,634 (5.06)	2,042 (5.18)	4,866 (5.61)	2,572 (4.44)
TOTAL	44,190 (100.00)	74,855 (100.00)	85,720 (100.00)	32,269 (100.00)	39,430 (100.00)	86,680 (100.00)	57,912 (100.00)
Average Number of Automobiles per Household	1.30	1.26	1.36	1.27	1.30	1.30	1.18

(Numbers in brackets are percentages)

Source: U. S. Bureau of Census, 1970.

In order to test whether this distribution is independent of the urban area (i.e., similar for all areas), a Chi-Square test similar to the one used in the previous chapter and described in the Appendix can be used. A 7×4 contingency table can be developed with "Urban Area" as the row factor (7 different levels) and "Auto Ownership" as the column factor (4 different levels). The frequencies observed for each combination of factor levels consist of the number of households in each auto ownership group in each of the urban areas, as observed in the original census sample survey. In other words, the frequencies used for testing purposes should not be the numbers reported in Table 17 because these have been expanded by the Bureau of Census from the raw sample survey data to characterize all the households in the SMSA. The expansion factor is a function of the sampling rate used for the survey, which in this case is 20 percent. Therefore, the frequencies entered in the contingency table can be obtained by calculating 20 percent of the corresponding numbers shown in Table 17.

The contingency table and the corresponding Chi-Square test are summarized in Table 18. This table shows that the hypothesis of independence between household distribution by auto ownership and "urban area" is rejected with 99.5 percent confidence. Since the differences in the distribution of auto ownership between urban areas cannot be attributed to sampling error alone, it is concluded that the distribution of households among auto ownership classes is not the same for all these urban areas.

The above test is a global test of independence for all the urban areas taken together. It does not exclude similarity between pairs or

Table 18. Summary of Chi-Square Test for Household
Distribution by Auto Ownership.

<u>Urban Area</u>	<u>Auto Ownership</u>			
	<u>0</u>	<u>1</u>	<u>2</u>	<u>3⁺</u>
ANDERSON	970	4,722	2,733	413
EVANSVILLE	2,340	7,168	4,710	753
FORT WAYNE	1,896	8,188	5,996	1,064
LAFAYETTE	720	3,598	1,809	327
MUNCIE	904	4,076	2,497	408
SOUTH BEND	2,123	8,926	5,313	973
TERRE HAUTE	1,922	6,140	3,006	514

H_0 : household distribution by auto ownership is independent
of the urban area

H_1 : household distribution by auto ownership is different in
each of the urban areas.

$\alpha = 0.005$ d.f. = 18

Test Statistic $\chi^2_{\text{data}} = 685.63$

$$\chi^2_{18,0.005} = 37.156 < \chi^2_{\text{data}}$$

Conclusion: Reject H_0 .

groups of urban areas. For this reason, a finer analysis was undertaken. It consisted of pairwise comparisons of these urban areas. Chi-Square tests on 2×4 contingency tables (2 urban areas and 4 levels of auto ownership) were run. These are summarized in Table 19.

The hypothesis tested was that household distribution by auto ownership was similar in the areas compared, or, in statistical terms, that the two samples could have come from the same population. This hypothesis was rejected for 19 out of the 21 comparisons with 99.5 percent confidence. It was also rejected for one other comparison with 90 percent confidence (Comparison of Muncie and South Bend). It was accepted in one case only: comparison of Anderson and Muncie.

The decision to reject the hypothesis in the test of similarity between Muncie and South Bend was made after very careful evaluation. The problem amounted to deciding on the acceptable risk of making a Type I error (rejecting the hypothesis when it is actually true). In this comparison, while it was not possible to reject the hypothesis with 95 percent confidence, it was possible to reject it with 90 percent confidence. Judgement had to be exercised in this decision because it was not clear cut like most of the others. This judgement was based on a careful evaluation of the context in which the Chi-Square test was applied. Since the auto-ownership data was derived from census data relative to the whole SMSA's, the results of the test would apply to the SMSA's being compared. However, while the unit of analysis in this study being the urban area as defined for the transportation planning studies, it was felt that the results obtained by comparing SMSA's would hold for the transportation study areas as well. Moreover, it

Table 19. Summary of Chi-Square Tests for Pairwise Comparisons of Household Distribution by Auto Ownership

<u>Urban Areas Compared</u>	<u>χ^2_{data}</u>	<u>d.f.</u>	<u>α</u>	<u>$\chi^2_{\text{d.f.,}\alpha}$</u>	<u>Decision</u>
Anderson & Evansville	123.00	3	0.005	12.838	Reject H_0
Anderson & Fort Wayne	90.00	3	0.005	12.838	Reject H_0
Anderson & Lafayette	15.60	3	0.005	12.838	Reject H_0
Anderson & Muncie	6.15	3	0.1	6.251	Accept H_0
Anderson & South Bend	22.00	3	0.005	12.838	Reject H_0
Anderson & Terre Haute	>50.00	3	0.005	12.838	Reject H_0
Evansville & Fort Wayne	>75.00	3	0.005	12.838	Reject H_0
Evansville & Lafayette	136.00	3	0.005	12.838	Reject H_0
Evansville & Muncie	>60.00	3	0.005	12.838	Reject H_0
Evansville & South Bend	>45.00	3	0.005	12.838	Reject H_0
Evansville & Terre Haute	111.00	3	0.005	12.838	Reject H_0
Fort Wayne & Lafayette	139.00	3	0.005	12.838	Reject H_0
Fort Wayne & Muncie	>65.00	3	0.005	12.838	Reject H_0
Fort Wayne & South Bend	>45.00	3	0.005	12.838	Reject H_0
Fort Wayne & Terre Haute	>95.00	3	0.005	12.838	Reject H_0
Lafayette & Muncie	26.92	3	0.005	12.838	Reject H_0
Lafayette & South Bend	34.20	3	0.005	12.838	Reject H_0
Lafayette & Terre Haute	>60.00	3	0.005	12.838	Reject H_0
Muncie & South Bend	6.50	3	0.1	6.251	Reject H_0
Muncie & Terre Haute	>100.00	3	0.005	12.838	Reject H_0
South Bend & Terre Haute	> 60.00	3	0.005	12.838	Reject H_0

H_0 : Household Distribution by Auto Ownership is independent of the urban area, i.e., it is similar in the areas compared.

was noticed that the cordon area for the South Bend transportation study was considerably smaller than the corresponding SMSA, the fact that called for even more precaution. Rejecting the hypothesis was considered to the "safe" and conservative decision, especially since the hypothesis was rejected in most of the other tests. In addition, considering that Anderson and Muncie were shown to be similar, it would have been intuitively (even though not formally and statistically) expected that Anderson and South Bend would be similar if Muncie and South Bend were similar. However, the Anderson and South Bend comparison led to the rejection of the hypothesis of similarity. Hence, and for all of the above considerations, the decision was to reject the hypothesis of the similarity of the household auto ownership distribution in Muncie and South Bend.

The above tests lead to the conclusion that only two of the urban areas, Anderson and Muncie, have similar household distributions by auto ownership. All the others are different with respect to each other. The implications of these results on trip frequency transferability will be discussed in the next section.

Implications for Trip Frequency Transferability and Synthetic Applications

The socio-economic similarity between Anderson and Muncie is also reflected in the areawide trip frequency rates obtained from these two urban areas. As shown in Table 9, the areawide average household vehicle trip rates are 9.74 and 8.92 for Anderson and Muncie, respectively. Moreover, as shown in Table 12, the areawide average household person trip rates are 13.18 and 12.95 for Anderson and Muncie

respectively. Both vehicle and person trip rates are similar for the two urban areas.

Areawide household trip frequency parameters therefore depend on the socio-economic mix of the households in the urban area under study. Thus, in order to transfer aggregate areawide parameters from one urban area to another, care must be taken to thoroughly compare the socio-economic characteristics of the households and residents of those two areas. In most cases, these characteristics would be different between one area and another, at least from a strict statistical standpoint. This raises the issue discussed in earlier chapters concerning the trade off between the accuracy of the analysis tools and modeling techniques used in the planning process on one hand and the cost of such accuracy and sophistication on the other hand.

For gross areawide estimations, transferring aggregate trip frequency parameters between urban areas is certainly feasible, especially if the socio-economic characteristics of these areas are judged to be roughly similar (without going through detailed statistical analysis). However, more accurate areawide predictions would require more detailed and careful comparison of the socio-economic characteristics of the areas from and to which parameters are being transferred. More detailed analysis and forecasting at a finer level of aggregation would most definitely not rely on borrowed aggregate areawide travel frequency parameters but should try to borrow parameters developed at a finer level of aggregation, such as the traffic zone or the individual household. Trip frequency at these finer levels of aggregation is examined in the following chapters.

A final note must be made about the fact that auto ownership has been used in the tests conducted in this chapter only as an indicator of socio-economic characteristics. This does not mean that it is the only important factor in determining areawide travel frequency. The finer the categorization of the households by various socio-economic characteristics in urban areas, the more accurate is the comparison between these areas, and the more reliable are results obtained using borrowed parameters.

CHAPTER 5: TRIP GENERATION ANALYSIS AT THE ZONAL LEVEL

Most transportation studies use the zone as the basic geographic unit of analysis. Multiple regression techniques are generally used to relate internal zonal trips (or average zonal trip rates) to the zonal socio-economic and land use characteristics. Zonal trip generation regression analysis is a standard and very familiar technique which will not be discussed in this chapter. Only a brief summary of this technique will be presented, in order to clarify some of the terminology used throughout this chapter. Regression equations developed in each of the study areas will be presented, and their transferability will be investigated through their predictive ability in areas other than the ones where they were developed. These equations will be critically evaluated and recommendations for synthetic modeling applications will be made.

Regression Technique

The use of linear regression for trip generation analysis is very thoroughly documented in the literature. The statistical aspect of the regression technique is discussed in many standard textbooks on the subject, such as references 26 and 35. Specific applications to trip generation can be found in most textbooks on transportation planning models (e.g., References 36, 37 and 38) as well as various other documents, such as Reference 39.

The object of the linear regression technique is to relate quantitatively zonal productions and attractions (dependent variables) to the

zonal socio-economic and land use characteristics (independent variables). Linear regression equations usually having the following form:

$$Y = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n$$

where

Y = dependent variable, which is the number (or rate)
of trips produced or attracted in a zone.

a_0 = a constant

$X_1, X_2, \dots, X_n$ = set of independent variables: zonal socio-economic
and land use characteristics

$a_1, a_2, \dots, a_n$ = coefficients of the independent variables.

Calibration of the regression equations at the zonal level requires a set of observations which consists of the values for both the dependent and the independent variables for each of the zones of an urban area. The coefficients $a_0, a_1, \dots, a_n$ are thus determined to best fit these observations.

Review of Past Synthetic Efforts Using Zonal Regression Equations

A feeling exists among many planners that zonal regression equations developed in a certain area could be transferred and used to model travel in a different area, especially in the case of small urban areas (16,19, 20,40). Equations based on pooled data have also been developed specifically for synthesizing travel patterns in small urban areas (19,40). A summary of these efforts is presented hereafter.

The Georgia State Department of Highways has developed a set of trip generation equations for zonal productions and attractions based on equations developed in four urban areas with populations in the range

of 13,000 to 50,000 (19). These equations are:

Productions

$$\text{Home Based Work: } 8 + 1.2 \times (\text{Autos})$$

$$\text{Home Based Other: } 18 + 2.7 \times (\text{Autos})$$

$$\begin{aligned} \text{Non-Home Based: } & 67 + 0.1 \times (\text{Dwelling Units}) + 0.1 \times (\text{Autos}) \\ & + 0.6 \times (\text{Employment}) + 0.1 \times (\text{School} \\ & \text{Enrollment}) \end{aligned}$$

$$\begin{aligned} \text{Total Internal Productions: } & 164 + 4.2 \times (\text{Autos}) + 0.4 \times \\ & (\text{Employment}) \end{aligned}$$

Attractions

$$\text{Home Based Work: } 54 + 0.9 \times (\text{Employment})$$

$$\begin{aligned} \text{Home Based Other: } & 206 + 0.3 \times (\text{Dwelling Units}) + 0.6 \times \\ & (\text{Employment}) + 0.3 \times (\text{School Enrollment}) \end{aligned}$$

$$\begin{aligned} \text{Non-Home Based: } & 67 + 0.5 \times (\text{Dwelling Units}) + 0.4 \times \\ & (\text{Employment}) \end{aligned}$$

$$\text{Total Attractions: } 221 + 2.3 \times (\text{Autos}) + 2.3 \times (\text{Employment}).$$

These equations were evaluated in four other urban areas of the same size group in Georgia, and the tests led to their acceptance for use to synthesize travel for transportation studies in small urban areas. It was also noted that since attraction equations were less reliable than those for productions, the standard practice of "balancing" attractions to productions should be maintained for studies based on those equations.

The Kentucky Highway Department also has developed a set of synthetic trip generation equations (40). These are based on areawide trip frequency rates, areawide purpose distribution of trips, and the zonal

share of areawide totals of socio-economic and land use characteristics. These equations are:

Productions

Home Based Work: $0.40 \times (\text{zonal population})$

Home Based Other: $0.96 \times (\text{zonal population})$

Non-Home Based: $0.64 \times (\text{total population}) \times$
 $\left(\frac{\text{zonal employment}}{\text{total employment}} \right)$

Attractions

Home Based Work: $0.40 \times (\text{total population}) \times$
 $\left(\frac{\text{zonal employment}}{\text{total employment}} \right)$

Home Based Other: $0.96 \times (\text{total population}) \times$
 $\left(\frac{\text{zonal non-industrial employment}}{\text{total non-industrial employment}} \right)$

Non-Home Based: $0.64 \times (\text{total population}) \times$
 $\left(\frac{\text{zonal employment}}{\text{total employment}} \right)$

This set of equations has the advantage of using only a few independent variables, which reduces data collection and forecasting requirements.

The equations from Kentucky and Georgia as well as another set of equations developed in a particular urban area were comparatively tested for their ability to synthesize travel in three small urban areas in Tennessee (16). Two types of comparisons were made:

- synthesized productions and attractions were compared to those observed in the O-D survey previously collected in the test areas.
- assigned link volumes obtained after distributing the synthetically derived productions and attractions were compared to actual volume ground counts.

It was concluded from these tests that a great deal of variation occurred between the results obtained using the various synthetic models. Elaborate verification checks were recommended if synthetic equations were to be used. The importance of accurate control totals was also stressed. As to the "best" set of synthetic equations among those tested, it was concluded that "fabricated models based upon pooled data over a range of population sizes, such as the Georgia equations provided estimates of zonal productions and attractions consistent with the more stratified and detailed models" (16).

Another set of synthetic zonal trip generation equations was developed by Wilson and Kristoffersen (20). These were developed using both dwelling units and autos as independent variables in total trip generation equations and then applying a step down procedure based on fixed areawide percentages of trips by purpose. Home Based Work, Home Based Other and Non-Home Based productions equations were thus developed, as well as an equation for Home Based Work attractions. Non-linear trip generation equations were developed as well. Eighteen urban areas with populations of less than 100,000 served as test areas for the predictive accuracy of these synthetic equations. Special emphasis was placed on Home Based Work productions for which it was found that the synthesized results compared favorably with mean observed values, with the best overall performance observed by using dwelling units as the independent variable. It was also found that the non-linear Home Based Work equation developed led to more accurate productions estimates for small sized zones and resulted in smaller standard errors.

The potential transferability of the internal vehicle trip generation equations developed in the Indiana urban areas considered in this

study will be investigated in the following sections.

Internal Vehicle Trip Generation Equations Developed in Study Areas

All seven study areas have developed zonal trip generation equations. The trip purposes for which these equations were developed varied between the study areas; however, these equations have been combined appropriately according to the three purposes defined for this study in Chapter 3 (HBW, HBO and NHB) in order to have a basis for comparison. For example, if a certain area had developed three separate equations for Home Based Shopping, Home Based Social and Home Based Miscellaneous trips, these three equations were combined into one equation for Home Based Other trips. This was done on an additive basis, the rationale behind it being that in order to obtain predictions for HBO trips for that area, the separate predictions for Social, Shopping and Miscellaneous trips would have had to be added. The equation thus obtained however might not be the one that would have been obtained had it been fitted directly to Home Based Other trips.

The equations reported here are those that have been adopted by the study areas for the development of their future travel forecasts. These equations are therefore not necessarily the best statistically, because other factors have to be taken into account when selecting final trip generation equations, such as the number of independent variables in the equation (as few as possible), data availability and the ease of forecasting of these independent variables.

Trip generation equations for internal vehicle trip Home Based Work productions and Home Based Other productions for each of the study areas are shown in Table 20. Non-Home Based productions and attractions

Table 20. Internal Vehicle Trip Production Regression Equations
Developed in Study Areas. (HBW and HBO)

Home Based Work Productions (HBW)

<u>Urban Area</u>	<u>Coefficients</u>				
	<u>Constant</u>	<u>Population</u>	<u>Autos</u>	<u>Employed Residents</u>	<u>Students</u>
ANDEPSON	-2	---	---	1.37	---
FORT WAYNE	-15.6531	---	1.2383	---	---
LAFAYETTE	9.894	0.4949	---	---	---
MUNCIE	1.1288	0.4159	---	---	---
SOUTH BEND	---	---	1.0263	---	---
TERRE HAUTE	8.14258	---	---	1.02462	---

Home Based Other Productions (HBO)

ANDERSON	38	---	2.3	---	---
FORT WAYNE	-38.4014	---	2.0514	---	---
LAFAYETTE	1.144	1.5747	---	---	---
MUNCIE	-48.6121	---	4.3008	---	---
SOUTH BEND	---	---	1.98732	---	0.06929
TERRE HAUTE	98.7102	---	2.78088	---	---

equations are shown in Table 21, while Home Based Work and Home Based Other attractions equations are shown in Table 22. It can be noticed that no equations are reported for Evansville. This is due to the fact that the trip generation equations developed in Evansville were for person trips and not for vehicle trips. These person trip generation equations are nevertheless shown in Table 23.

The ability of each of the vehicle trip generation equations developed in each of the study areas to reproduce the internal vehicle trip productions and attractions in the other study areas is investigated hereafter.

Transferability of the Internal Vehicle Trip Production Equations

The transferability of the internal vehicle trip production equations will be investigated at two levels of detail:

1. For each urban area, the mean zonal value of trips produced for each purpose will be predicted using the equations calibrated in other areas. Areawide mean zonal values of the independent variables will be used for these predictions. The predicted values thus obtained will be compared to the corresponding observed values (from the O-D survey). One measure of the overall predictive ability of each equation in each urban area will be calculated; it is the percent difference of the mean (difference between mean of estimated zonal trips and mean of observed zonal trips as a percent of the observed mean).

2. A more detailed evaluation will be made for selected urban areas (based on data availability) by predicting zonal trips for each purpose using zonal values of the independent variables. These predictions will be compared on a zone by zone basis to the corresponding

Table 21. Non-Home Based Internal Vehicle Trip Production and
Attraction Equations Developed in Study Areas

Non-Home Based Productions		Coefficients							
Urban Area	Constant	Employed Residents	Occupied Dwelling Units	Population	Autos	Total Employment	Manufacturing Employment	Retail Employment	Other Employment
ANDERSON	51	---	---	---	0.64	0.31	---	3.74	---
FORT WAYNE	29.6181	---	0.9911	0.0193	---	0.3472	---	2.6195	---
LAFAYETTE	94.951	---	---	0.2841	---	---	---	3.8444	0.5933
MUNCIE	26.1385	---	---	---	0.6778	---	0.2737	3.18835	0.6962
SOUTH BEND*									
TERRE HAUTE	64.70694	0.29883	---	---	0.61788	0.58035	---	---	---
Non-Home Based Attractions									
ANDERSON	51	---	---	---	0.64	0.31	---	3.74	---
FORT WAYNE	41.2713	---	1.0689	0.0210	---	0.2328	---	2.8433	---
LAFAYETTE	125.985	---	---	0.2286	---	---	---	4.2205	0.4364
MUNCIE	47.4992	---	---	0.2697	---	---	0.1612	3.3013	0.5701
SOUTH BEND*									
TERRE HAUTE	87.26828	0.25259	---	0.18096	---	0.66749	---	---	---

*South Bend NHB equations are non-linear.

Table 22. Internal Vehicle Trip Attraction Equations
Developed in Study Areas (HBW and HB0)

Home Based Work Attractions (HBW)

Urban Area	Constant	Employed Residents	Units Dwelling Units	Population	Autos	Coefficients			
						Total Employment	Manufacturing Employment	Retail Employment	Other Employment
ANDERSON	80	---	---	---	---	0.62	---	---	---
FORT WAYNE	101.5332	---	0.2774	---	---	0.5608	---	---	---
LAFAYETTE	68.061	---	---	---	---	---	0.6999	0.9775	0.7320
MUNCIE	- 2.8746	---	---	---	---	0.7025	---	---	---
SOUTH BEND	---	---	---	---	---	1.09211	---	---	---
TERRE HAUTE	82.59061	---	---	---	---	0.57707	---	---	---

Home Based Other Attractions (HB0)

ANDERSON	178	---	---	---	---	1.60	-1.36	3.60	---
FORT WAYNE	94.0329	---	1.1619	---	---	---	---	3.8057	---
LAFAYETTE	123.267	---	---	0.4997	---	---	---	5.3633	---
MUNCIE	66.0234	---	---	0.3118	---	---	0.1477	3.5818	1.2416
SOUTH BEND*									
TERRE HAUTE	99.38044	---	---	0.47355	---	---	---	2.29001	---

*South Bend HB0 equation is non-linear.

Table 23. Internal Person Trip Generation Equations - EvansvillePRODUCTIONS

$$\text{Home Based Work: } Y = -17.15 + 0.56 \times (\text{Population})$$

$$\text{Home Based Other: } Y = 25.50 + 4.14 \times (\text{Auto}) \quad \text{Urban*}$$

$$Y = 14.67 + 2.77 \times (\text{Auto}) \quad \text{Rural*}$$

$$\begin{aligned} \text{Non-Home Based: } Y = & 47.62 + 4.43 \times (\text{Retail Employment}) + \\ & 0.96 \times (\text{Institution Employment}) \\ & + 0.55 \times (\text{Autos}) + 0.27 \times (\text{Total} \\ & \text{Employment}) \quad \text{Urban} \end{aligned}$$

$$\begin{aligned} Y = & -5.55 + 1.75 \times (\text{Institution Employment}) \\ & + 0.76 \times (\text{Dwelling Units}) + 0.22 \times (\text{Total} \\ & \text{Employment}) \quad \text{Rural} \end{aligned}$$

ATTRACTIONS

$$\text{Home Based Work: } Y = 35.86 + 1.09 \times (\text{Total Employment})$$

Home Based Other:

$$Y = 132.98 + 1.56 \times (\text{School Enrollment}) + 3.55 \times (\text{Retail Employment}) + 0.77 \times (\text{Autos}) \quad \text{Urban}$$

$$Y = 32.60 + 0.61 \times (\text{School Enrollment}) + 0.28 \times (\text{Population}) + 0.52 \times (\text{Total Employment}) \quad \text{Rural}$$

Non-Home Based:

$$\begin{aligned} Y = & 17.37 + 4.75 \times (\text{Retail Employment}) + 0.65 \times (\text{Autos}) \\ & + 0.80 \times (\text{Institution Employment}) + 0.20 \times (\text{Total} \\ & \text{Employment}) \quad \text{Urban} \end{aligned}$$

$$Y = -1.79 + 1.30 \times (\text{Institution Employment}) + 0.70 \times (\text{Dwelling Units}) + 0.27 \times (\text{Total Employment}) \quad \text{Rural}$$

*Evansville had separate equations for Rural and Urban zones.

O-D observations. Another measure of the overall predictive ability besides the percent difference of the mean can thus be calculated. It is the percent standard error of the estimate, expressed as a percentage of the mean value, which is a measure of the accuracy of a value predicted using the regression equation. A value below 25 percent would be excellent, below 50 percent good, but above 100 percent would be no good.

These comparisons will be discussed below for each of the three trip purposes:

a. Home Based Work Productions

The analysis at the first level of detail (as described above) is summarized in Table 24. This table shows the comparisons between the observed mean zonal trips for each of the study areas and the mean zonal trips as predicted by the equations reported in Table 20. Overall, most of the predicted values are within 20 percent of the observed values. Notable exceptions are values predicted using the equation calibrated in Anderson which had "employed residents" as the independent variable. The most consistently adequate performance is observed for the equation developed in South Bend ($HBW = 1.0263 \times \text{Autos}$).

This observation is confirmed by the more detailed zone by zone comparison undertaken for three of the study areas where reliable zonal data was readily available: Fort Wayne, Muncie and Lafayette. The results of these comparisons are summarized in Table 25. The percent standard error of the estimate is acceptable for all the cases shown. The South Bend equation performs exceptionally well in all three study areas in terms of that statistic. This equation will be singled out for

Table 24. Comparison of Estimated Mean Zonal Trips to Observed Mean Zonal Trips -

Home Based Work Productions								
Urban Area		ANDERSON	EVANSVILLE	FORT WAYNE	LAFAYETTE	MUNCIE	SOUTH BEND	TERRE HAUTE
Equation								
ANDERSON	Obs. Mean	324.8	255.4	482.3	445.3	205.1	514.0	322.3
	Est. Mean	325.8	291.8	n.a.	435.8	264.8	678.1	417.5
	Difference	1	36.4		-9.5	59.7	164.1	95.2
	% Difference	0.3	14.3		-2.1	29.1	31.9	29.5
FORT WAYNE	Est. Mean	312.8	318.3	483.0	488.8	235.5	604.3	428.3
	Difference	-12	62.9	0.7	43.5	30.4	90.2	106.0
	% Difference	-3.7	24.6	0.1	9.8	14.8	17.6	32.9
LAFAYETTE	Est. Mean	313.5	307.4	546.8	426.9	248.2	629.3	452.8
	Difference	-11.3	52.0	64.5	-18.4	43.1	115.3	130.5
	% Difference	-3.5	20.4	13.4	-4.1	21.0	22.4	40.5
MUNCIE	Est. Mean	256.3	251.1	452.4	351.5	201.4	521.6	373.3
	Difference	-68.5	-4.3	-29.9	-93.8	-3.7	7.6	51.0
	% Difference	-21.1	-1.7	-6.2	-21.0	-1.8	1.5	15.8
SOUTH BEND	Est. Mean	272.2	276.8	412.7	418.1	208.0	513.8	368.0
	Difference	-52.6	21.4	-69.6	-27.2	2.9	-0.2	45.7
	% Difference	-16.2	8.4	-14.4	-6.1	1.4	0.0	14.2
TERRE HAUTE	Est. Mean	253.4	227.8	519.4	335.6	206.2	516.8	321.9
	Difference	-71.4	-27.6	37.1	-109.7	1.1	2.8	-0.4
	% Difference	-22.0	-10.8	7.7	-24.6	0.5	0.6	-0.1

Table 25. Summary of the Predictive Ability of HBW Vehicle Trip

Origin of Equation	Production Equations Developed in Study Areas				
	Observed Mean	Estimated Mean	Difference	Percent Difference	SE of Estimate
FORT WAYNE ZONAL DATA					
FORT WAYNE	482.31	483.05	+0.74	+0.15	104.9
LAFAYETTE	482.31	546.829	+64.519	+13.38	256.7
MUNCIE	482.31	452.354	-29.957	-6.21	232.5
SOUTH BEND	482.31	412.693	-69.617	-14.43	153.4
LAFAYETTE ZONAL DATA					
ANDERSON	445.276	435.833	-9.443	-2.12	97.87
FORT WAYNE	445.276	488.771	+43.495	+9.77	128.5
LAFAYETTE	445.276	426.85	-18.426	-4.14	108.38
MUNCIE	445.276	351.527	-93.749	-21.05	160.08
SOUTH BEND	445.276	418.058	-27.218	-6.11	120.68
TERRE HAUTE	445.276	335.597	-109.679	-24.63	177.43
MUNCIE ZONAL DATA					
FORT WAYNE	205.097	235.503	+30.406	+14.8	82.4
LAFAYETTE	205.097	248.236	+43.139	+21.03	99.37
MUNCIE	205.097	201.425	-3.672	-1.79	80.98
SOUTH BEND	205.097	207.991	+2.894	+1.41	58.70
					40.2
					48.4
					39.5
					28.6

synthetic applications, and is recommended for use in other areas, especially since its independent variable (autos in zone) is usually readily available and easy to forecast.

b. Home Based Other Productions

The same comparisons as those made for HBW productions are shown for HBO productions in Table 26 and 27. It is very obvious from these results that borrowing HBO productions regression equations developed in other areas does not lead to satisfactory results. Drastically erroneous estimates were obtained in many instances, especially when equations developed in areas with high areawide HBO trip frequency rates (cf. Tables 9-11) were used to estimate zonal trips in areas with lower areawide rates, or vice-versa. For example, the equations developed in Lafayette and Muncie (which have high areawide trip rates) consistently overpredicted the mean values in all of the other areas. The use of synthetic HBO productions calibrated at the zonal level is therefore strongly discouraged.

c. Non-Home Based Productions

The comparisons of estimated mean zonal trips to the observed values are shown in Table 28. This table shows that estimated values compare fairly well with observed values in most of the study areas with one notable exception: South Bend. All of the equations underpredicted by about 40 percent the mean zonal non-home based vehicle trips in South Bend. This can be attributed to the fact that South Bend, being larger than the other areas, has a very large number of truck trips which have been lumped with NHB trips (as explained in Chapter 3). Because of the relatively large number of truck trips in South Bend, including these trips with NHB trips might not be justified for that

Table 26. Comparison of Estimated Mean Zonal Trips to Observed Mean Zonal Trips -
Home Based Other Productions

Equation	Urban Area		ANDERSON	EVANSVILLE	FORT WAYNE	LAFAYETTE	MUNCIE	SOUTH BEND	TERRE HAUTE
	Obs.	Mean							
ANDERSON	Est.	Mean	718.7	658.4	962.8	1018.3	470.9	1227.6	831.3
HB0	Difference		-13.1	13.6	176.8	-380.8	-295.9	36.4	-184.0
EQUATION	% Difference		-1.8	2.1	22.5	-27.2	-38.6	3.1	-18.1
FORT WAYNE	Est.	Mean	568.7	514.9	798.2	836.5	350.0	1022.6	669.1
HB0	Difference		-163.1	-129.9	3.2	-562.6	-416.8	-168.5	-346.2
EQUATION	% Difference		-22.3	20.1	0.4	-40.2	-54.4	-14.1	-34.1
LAFAYETTE	Est.	Mean	982.2	947.7	1710.0	1399.1	706.9	1971.9	1349.2
HB0	Difference		250.4	302.9	924.0	0	-59.9	780.8	333.9
EQUATION	% Difference		34.2	47.0	117.6	0	-7.8	65.5	32.9
MUNCIE	Est.	Mean	1092.1	1111.4	1683.0	1784.8	762.4	2175.8	1434.7
HB0	Difference		360.3	466.6	897.0	385.6	-4.4	984.7	419.4
EQUATION	% Difference		49.2	72.4	114.1	27.6	-0.6	82.7	41.3
SOUTH BEND	Est.	Mean	600.5	548.6	n.a.	853.6	394.0	1054.7	708.2
HB0	Difference		-131.3	-96.2		-545.5	-372.8	-136.4	-307.1
EQUATION	% Difference		-17.9	-14.9		-39.0	-48.6	-11.5	-30.2
TERRE HAUTE	Est.	Mean	921.7	848.8	1216.8	1283.9	622.1	1537.0	1057.8
HB0	Difference		189.9	204.0	430.8	-115.2	-144.7	345.9	42.5
EQUATION	% Difference		-26.0	31.6	54.8	-8.2	-18.9	29.0	4.2

Table 27. Summary of the Predictive Ability of HB0 Vehicle Trip
Production Equations Developed in Study Areas

Origin of Equation	Observed Mean	Estimated Mean	Difference	Percent Difference	SE of Estimate	Percent SE of Estimate
FORT WAYNE ZONAL DATA						
ANDERSON	785.914	962.772	+176.858	+22.5	258.5	32.88
FORT WAYNE	785.914	789.165	+3.251	+0.414	157.02	19.98
LAFAYETTE	785.914	1710.03	+924.116	+117.6	1431.7	182.2
MUNCIE	785.914	1682.983	+897.069	+114.14	1299.3	165.3
TERRE HAUTE	785.914	1216.832	+430.918	+54.8	549.7	69.9
LAFAYETTE ZONAL DATA						
ANDERSON	1399.137	1018.26	-380.877	-27.2	691.7	49.43
FORT WAYNE	1399.137	836.529	-562.608	-40.21	856.1	61.19
LAFAYETTE	1399.137	1399.08	-0.057	-0.004	522.4	37.34
MUNCIE	1399.137	1784.765	+385.628	+27.6	612.2	43.76
TERRE HAUTE	1399.137	1283.921	-115.216	-8.23	482.4	34.48
MUNCIE ZONAL DATA						
ANDERSON	766.807	470.855	-295.951	-38.6	530.75	69.2
FORT WAYNE	766.807	349.981	-416.826	-54.36	639.57	83.41
LAFAYETTE	766.807	706.937	-59.87	-7.81	353.6	46.12
MUNCIE	766.807	762.36	-4.447	-0.58	220.5	28.75
TERRE HAUTE	766.807	622.066	-144.741	-18.88	390.64	50.9

Table 28. Comparison of Estimated Mean Zonal Trips to Observed Mean Zonal Trips -
Non-Home Based Productions

Equation	Urban Area		EVANSVILLE	FORT WAYNE	LAFAYETTE	MUNCIE	SOUTH BEND	TERRE HAUTE
	Obs. Mean	Est. Mean						
ANDERSON			647.0	765.5	744.6	300.6	1382.3	518.1
NHB			524.3	711.6	722.6	324.9	853.0	597.5
EQUATIONS			-110.9	-53.9	-21.6	24.3	-529.3	79.4
	% Difference		-17.1	-7.0	-2.9	8.1	-38.3	15.3
FORT WAYNE			491.0	794.9	659.5	303.5	833.6	621.0
NHB			-156.0	29.4	-85.1	2.9	-548.7	102.9
EQUATION			-24.1	3.8	-11.4	1.0	-39.7	19.9
	% Difference							
LAFAYETTE			557.4	783.8	743.5	299.9	889.2	663.2
NHB			-89.6	18.3	-1.0	-0.7	-493.1	145.1
EQUATION			-13.8	2.4	-0.1	-0.2	-35.7	28.0
	% Difference							
MUNCIE			497.8	709.2	707.3	298.1	827.9	577.0
NHB			-149.2	-56.3	-37.3	-2.5	-554.4	58.9
EQUATION			-23.1	-7.4	-5.0	-0.8	-40.1	11.4
	% Difference							
TERRE HAUTE			459.5	720.7	628.6	285.8	825.2	490.5
NHB			-187.5	-44.8	-116.0	-14.8	-557.1	-27.6
EQUATION			-29.0	-5.9	-15.6	-4.9	-40.3	-5.3
	% Difference							

particular area. Another fact which is worthy of attention is that the NHB regression equations calibrated in South Bend contained non-linear terms, which indicates that the zonal NHB trip productions in that area require special consideration.

The unfavorable results obtained for South Bend should not rule out the use of synthetic Non-Home Based equations, especially since such equations will normally be used in areas smaller in size than South Bend which would be more comparable to Lafayette or Muncie where all synthetic equations performed well. A detailed zone by zone comparison for Lafayette is summarized in Table 29. No other area has been tested in detail for NHB trips because of high data requirements.

The results shown in Tables 28 and 29 substantiate the fact that NHB trip production equations have strong potential for transferability between smaller urban areas. All equations performed equally well in the test areas, with Lafayette's equation being one of the better ones

$$(NHB = 94.951 + 0.2841 \times \text{POPULATION} + 3.8444 \times \text{RETAIL EMPLOYMENT} + 0.5933 \times \text{OTHER EMPLOYMENT})$$

Transferability of the Internal Vehicle Trip Attraction Equations

The transferability of the internal vehicle trip attraction equations developed in the study areas will be investigated in less detail than that of the trip production equations. Various reasons exist for placing more emphasis on the production aspect of trip generation, the major one being that it is standard procedure to balance attractions to productions by constraining the total area wide attractions to equal the total areawide productions. This practice is based on the fact that residential trip production is "generally characterized by relationships

Table 29. Summary of the Predictive Ability of NHB Vehicle Trip Production Equations Developed in Study Areas.

Origin of Equation	LAFAYETTE ZONAL DATA				
	Observed Mean	Estimated Mean	Difference	Percent Difference	SE of Estimate
ANDERSON	744.579	722.588	-21.991	-2.95	333.0
FORT WAYNE	744.579	659.445	-85.134	-11.43	390.3
LAFAYETTE	744.579	743.526	-1.053	-0.14	308.9
MUNCIE	744.579	707.307	-37.272	-5.0	342.9
TERRE HAUTE	744.579	628.605	-115.974	-15.57	764.6
					102.7

that are more accurate and more stable"(8). Hence, many "unique" and non-residential generators such as hospitals, universities, shopping centers, etc... require special site-specific analyses because their travel patterns cannot be reproduced by generalized regression equations.

Moreover, various procedures for the estimation of trip attractions have been described in the literature (16, 22, 28, 41, 42). There also exists a very large data base of trip rates which have been compiled nationally and are being updated on a continuing basis (41,42). These procedures, nationwide rates, as well as special analyses conducted at the local level, can provide estimates of adequate accuracy for synthetic applications, especially since these estimates are to be adjusted according to the areawide control totals (to be established by more reliable techniques as discussed in the next chapter).

For the above reasons, the investigation of the transferability of the trip attraction regression equations will be limited to comparisons of the mean zonal trips (observed vs. estimated) for three of the study areas. These comparisons are shown in Tables 30, 31 and 32 for HBW, HBO and NHB trip attractions, respectively. Based upon these tables, the following observations can be made:

- All HBW equations developed in other areas considerably over-predict in Muncie while Muncie's HBW equation underpredicts in the other areas.
- HBO attraction equations are noticeably more stable than HBO production equations (cf. Tables 26 and 27), as evidenced by the more accurate estimations of the mean values.

Table 30. Comparison of Estimated Mean Zonal Trips to Observed
Mean Zonal Trips - Home Based Work Attractions

<u>Urban Area</u>		<u>FORT WAYNE</u>	<u>LAFAYETTE</u>	<u>MUNCIE</u>
	Observed Mean	460.2	372.7	181.7
ANDERSON	Est. Mean	375.1	323.0	234.2
HBW	Difference	-85.1	-49.7	52.5
EQUATION	% Difference	-18.5	-13.3	28.9
FORT WAYNE	Est. Mean	460.2	406.5	285.6
HBW	Difference	0.0	33.8	103.9
EQUATION	% Difference	0.0	9.1	57.2
LAFAYETTE	Est. Mean	427.5	373.6	261.3
HBW	Difference	-32.7	0.9	79.6
EQUATION	% Difference	-7.1	0.2	43.8
MUNCIE	Est. Mean	331.5	272.5	171.8
HBW	Difference	-128.7	-100.2	-9.9
EQUATION	% Difference	-28.0	-26.9	-5.4
SOUTH BEND	Est. Mean	519.9	428.1	271.6
HBW	Difference	59.7	55.4	89.9
EQUATION	% Difference	13.0	14.9	49.5
TERRE HAUTE	Est. Mean	357.3	308.8	226.1
HBW	Difference	-102.9	-63.9	44.4
EQUATION	% Difference	-22.4	-17.2	24.4

Table 31. Comparison of Estimated Mean Zonal Trips to Observed
Mean Zonal Trips - Home Based Other Attractions

<u>Equation</u>	<u>Urban Area</u>	<u>FORT WAYNE</u>	<u>LAFAYETTE</u>	<u>MUNCIE</u>
	Observed Mean	786.0	968.7	396.7
ANDERSON	Est. Mean	954.2	876.8	481.8
HBO	Difference	168.2	-91.9	85.1
EQUATION	% Difference	21.4	-9.5	21.4
FORT WAYNE	Est. Mean	781.3	748.8	362.5
HBO	Difference	-4.7	-219.9	-34.2
EQUATION	% Difference	-0.6	-22.7	-8.6
LAFAYETTE	Est. Mean	1069.3	968.4	497.4
HBO	Difference	283.3	-0.3	100.8
EQUATION	% Difference	36.0	-0.0	25.4
MUNCIE	Est. Mean	966.1	837.0	398.9
HBO	Difference	180.1	-131.7	2.2
EQUATION	% Difference	22.9	-13.6	0.6
TERRE HAUTE	Est. Mean	787.0	684.0	355.6
HBO	Difference	1.0	-284.7	-41.1
EQUATION	% Difference	0.0	-29.4	-10.4

Table 32. Comparison of Estimated Mean Zonal Trips to Observed
Mean Zonal Trips - Non-Home Based Attractions

<u>Urban Area</u>		<u>FORT WAYNE</u>	<u>LAFAYETTE</u>	<u>MUNCIE</u>
	Observed Mean	765.5	722.8	300.6
ANDERSON	Est. Mean	711.6	715.6	330.9
NHB	Difference	-53.9	-7.2	30.3
EQUATION	% Difference	-7.0	-1.0	10.1
FORT WAYNE	Est. Mean	730.2	694.1	311.1
NHB	Difference	-35.3	-28.7	10.5
EQUATION	% Difference	-4.6	-4.0	3.5
LAFAYETTE	Est. Mean	752.5	722.8	388.2
NHB	Difference	-13.0	0.0	87.6
EQUATION	% Difference	-1.7	0.0	29.1
MUNCIE	Est. Mean	710.8	649.1	311.5
NHB	Difference	-54.7	-73.7	10.9
EQUATION	% Difference	-7.1	-10.2	3.6
TERRE HAUTE	Est. Mean	713.5	567.1	296.4
NHB	Difference	-52.0	-155.7	-4.2
EQUATION	% Difference	-6.8	-21.5	-1.4

- The HBO attraction equation developed in Muncie seems to be the most appropriate for synthetic applications. It overpredicts in Fort Wayne (the equation of which consistently underpredicts in other areas) and overpredicts in Lafayette (the equation of which consistently overpredicts in other areas), within reasonable ranges though.

- NHB Attraction equations show good overall potential for transferability. No equation will be singled out for synthetic uses, considering that most equations, especially the areas from Anderson, Fort Wayne and Muncie, performed equally well in those limited tests.

Overall, trip attraction equations can be transferred for the purpose of developing rough estimates that would be refined later on in the process. At least one equation for each purpose can be recommended for such uses: Lafayette's HBW equation, Muncie's HBO equation, and either of the NHB equations of Anderson, Fort Wayne or Muncie. A critical evaluation of trip generation analysis using zonal regression equations is presented in the next section.

Critical Review of Aggregate Zonal Trip Generation

The results obtained in the previous sections tend to validate many of the concerns expressed regarding the use of regression equations calibrated at the zonal level (8,16,43,44). Zonal aggregate model parameters tend to be dependent on the particular zonal scheme for which the model was developed, mainly because of the constant term in the regression equation (8,16). They are therefore not readily transferable to a different zonal scheme in another urban area (or even in

the same area) unless this scheme is extremely similar in size and composition to the initial one. This fact was observed in the above comparisons: areas where the average number of zonal trips is sizeably smaller than in other areas (e.g., Muncie, HBW Attractions, Table 30) tended to have overpredicted means when equations from other areas of larger zonal size were used.

The disadvantages of aggregate zonal regression equations are of even greater concern in the case of home-based productions (HBW and HBO). These equations use zonal aggregate data, and assume that the zones are homogeneous in terms of the socio-economic characteristics of the households in the zone; in many instances, this might not really be the case (43). Moreover, aggregation to the zonal level tends to conceal a considerable amount of the variation between individual households within the zone, thus masking the causal nature of the explanatory variables. These models are therefore not behavioral in nature, and as a consequence, they are not sensitive to changes occurring at the basic decision level of tripmaking (44).

Considering the non-behavioral nature and the zonal bias of these regression models, the reservations expressed about their transferability between urban areas are justified. In order to overcome some of their disadvantages, it was suggested that regression equations with "trips per household" as the dependent variable be calibrated at the household level, thus considering the household as the basic unit of analysis (8). Trip generation rates at the household level are discussed in the next chapter.

CHAPTER 6: TRIP GENERATION ANALYSIS AT THE DISAGGREGATE HOUSEHOLD LEVEL

The previous chapter discussed some of the problems associated with trip generation regression equations calibrated at the zonal level, and the limitations concerning their transferability. The fact remains, however, that zonal values of trips made for each purpose constitute the required output from the trip generation phase in the 4-step conventional transportation planning process. Zonal regression equations directly provide these zonal trip values. However, these zonal values can alternatively be obtained by aggregating the number of trips made by each household in each zone. Models considering the household as the basic unit of analysis and calibrated using disaggregate household data, are examined in this chapter, with special emphasis being placed on trip generation by category analysis (cross classification). The problems associated with category analysis as well as their suggested solutions will be reviewed, and one of them, the determination of the levels of the independent variables, will be discussed in detail. Cross-classified trip rate tables will be developed for two of the study areas, Lafayette and Evansville, and statistically compared in order to assess their stability between urban areas and to evaluate their transferability in view of their use for synthetic modeling efforts in small urban areas.

Disaggregate Household Trip Generation Models

Within the framework of the established transportation planning process, two types of disaggregate household trip generation models can be used: regression models calibrated with household data, and cross-classification (category analysis) models. The latter stratify households according to their socio-economic characteristics, and provide estimates of the trip rates for each of the household categories. Both types of models have been shown to predict equally well over the full range of households and appear to be indistinguishable with respect to sample size sensitivity (45).

Both types of household disaggregate models have the following advantages over zonal aggregate models (16,22,44):

- more efficient and complete use of survey information since it is used prior to aggregation.
- savings in data required to calibrate models.
- independence from any geographic aggregation or zonal scheme, which makes these models adaptable to various planning contexts each requiring different aggregation units; this flexibility makes them valuable tools for smaller scale and corridor analysis, and prevents them from becoming obsolete if a modified zonal scheme is to be adopted.
- use of Census data (household socio-economic characteristics)
- greater causal validity due to their behavioral nature since household models represent the basic relations at the decision level.

- temporal stability associated with causal relationships.
- transferability between urban areas.

These last two points are somewhat controversial and will be discussed in greater detail later in this chapter.

In addition, category analysis offers the following advantages over the regression technique (16,22,46):

- ability to express non-linear relationships.
- distribution-free characteristic whereby no limiting (and sometimes inaccurate) assumptions as to the statistical distribution of the variables have to be made as in the case of regression analysis.
- ease of understanding and simplicity of application.
- ease of monitoring and updating using small disaggregate samples or even subjective judgement (as explained in Chapter 7).

These reasons make cross-classification models preferable to regression models, especially for simplified transportation planning for small urban areas. The procedural aspect of cross-classification analysis will be described in greater detail in the following section.

Cross-Classification Procedure

The cross-classification technique for household trip production analysis stratifies household trip rates according to the major household socio-economic characteristics affecting trip making behavior. Extensive empirical evidence has been presented to show that the two major determinants of household trip generation are auto ownership and household size (34). Households can therefore be grouped according to these two characteristics, and trip rates for each household category

can be determined. Some studies have used income as one of the classifying variables, either instead or in addition to car ownership or household size (22). Actually, the greater the number of differentiating variables (or independent variables), the finer is the categorization and the more accurate are trip rates. However, categorization according to auto ownership and household size provides adequate reliability, especially since auto ownership is very highly correlated with income. Another refinement could be made using "persons over five years of age" instead of household size (number of persons in household); however data for the total number of household members is more readily available and forecasted than "persons over five".

After the household differentiating variables are selected, households are categorized accordingly. Average trip rates for each of the household socio-economic categories (or cells) are then developed. In order to do so, trip rates (by trip purpose) for the individual households are computed and accumulated within each cell. The average number of trips made by households as well as the standard deviation are then computed for each household socio-economic group. These rates can be plotted and a smooth curve can be fitted to the observations.

Each household category therefore has an average trip rate associated with it. In order to determine zonal (or whatever unit of aggregation is desired) trip productions, the number of households of each group within each zone are estimated and multiplied by the appropriate trip rate and these products are summed (36):

$$P_{i(c)} = \sum h_{i(c)} \times r(c)$$

where $P_{i(c)}$ = the number of trips produced by zone i by
category c households

$h_{i(c)}$ = number of households in zone i in category c

$r(c)$ = average trip rate for category c households

Various questions can be raised during the development of trip rate tables. One of them concerns testing the statistical significance of the various explanatory variables. Some efforts have been made in that direction, such as the use of analysis of variance techniques (45, 46). However, in as much as the major determinants of household travel have been identified, detailed statistical analysis for each new application might not be needed. More important than complex and difficult statistical tests are reasonableness checks and logical relationships expressed in the cross-classification (22).

Another question concerns the number of levels or categories to be used for each of the independent variables. This matter will be treated in the next section.

Levels of the Independent Variables

One of the frequent problems encountered in developing cross-classification trip tables is the determination of the number of levels to use for each of the independent variables (i.e., the number of rows and columns of the trip rate table). A suggested method is to check the standard deviation of the trip rates in the cells corresponding to the highest levels of each independent variable (22). If the standard deviation is high, then one additional level might be justified in order to reduce the variability of the trip rates by further categorizing them into more homogeneous groups.

An inherent weakness of that method is that it requires the determination of a cut-off point as to when a standard deviation is "high" or not. This decision is usually left up to the planner and his "feel" for the data at hand.

It is felt that a stronger statistical basis is needed for the determination of the appropriate level of disaggregation of the independent variables. An alternative systematic procedure for this purpose is described hereafter.

Two considerations must be taken into account when selecting the number of levels:

1. The availability of observations in the cells corresponding to higher levels of the independent variables.

2. The stability of the trip rates when going from one level to a higher one, i.e., whether the rates for the two adjacent higher level cells are different or not.

The first of these considerations is due to the low frequency of households corresponding to the "extreme" socio-economic cells (households with 10 members and 6 cars are less common than households with 4 members and 2 cars). Therefore, the presence of an adequate number of such households in the sample survey might not be very likely.

The decision as to whether the number of households in a specific cell is adequate to provide enough reliability again appeals to one's sound engineering judgement. However, previous research has shown that approximately 30 observations per cell is an adequate cut-off point (16).

One way of handling insufficient observations in certain cells is selective sampling, whereby households in less frequent categories are sampled more heavily in order to obtain adequate representation.

Another is the combination of the insufficiently sampled highest level of the same independent variable. This latter measure is only expected since households in the extreme cells are so infrequent that "lumping" them with the adjacent one-step-lower level cell and using the resulting average rate will most probably not significantly affect the results at the time of aggregation (at the zonal level). Moreover, chances are that the rates for these higher level cells are not significantly different from the rates for the adjacent cells, as will be discussed in the following paragraphs.

The second consideration when selecting the level of disaggregation of the independent variables is the stability of the trip rates after a certain level for each of the independent variables. Trip rates are considered stable with respect to a certain independent variable when they reach steady state, that is when further categorization based on that independent variable does not lead to a significant change in the rates in the resulting cells.

A statistical technique that can be used to test the stability of trip rates is a series of pairwise comparisons of the mean values of the average trip rates in adjacent cells with the aid of t-tests. These tests do not require assumptions as to the distribution of the trip rates for households within a certain cell, and therefore are in line with the previously discussed distribution-free characteristic of the cross-classification technique. Indeed, it is implied by the Central Limit Theorem that the means (and sums) of samples of size n from even a non-normal population follow an approximately normal distribution for large values of n (usually for $n \geq 30$). The test as applied to the problem

at hand will be discussed in further detail in a later section.

The overall formalized technique will be best described through a detailed application to an actual example: a cross-classification table for total vehicle trip rates per households, with household size and car ownership as independent variables, based on household level data from the O-D sample survey conducted in the Greater Lafayette area, Indiana. The general idea is to first group the data in a larger set of cells, and to subsequently combine these cells into a final set which characterizes the richness of the data. Following is a step by step description of the analysis:

Step 1: The cross-classification table was developed, with the following levels of the independent variables:

CARS: 0, 1, 2, 3, 4⁺

SIZE: 1, 2, 3, 4, 5, 6⁺

This table is shown in Table 33. The following information is available for each cell: the average trip rate for households in the cell, the number of sampled households in that cell, as well as the standard deviation of the trip rates within the cell. The following notation will be adopted:

$\bar{t}_{ij}$: sample average trip rate for cells corresponding to households owning i cars and of size j .

n_{ij} : number of sampled households in cell i, j

s_{ij} : standard deviation of trip rates within cell i, j .

Step 2: The number of observations in the cells corresponding to the highest levels of each of the independent variables were checked. All but one of the cells corresponding to the "4⁺" car ownership category

Table 33. Cross-Classification of Total Vehicle Trip Rates, LAFAYETTE
5 Levels of Auto Ownership and 6 Levels of Household Size

Autos Per Household	Household Size					
	1	2	3	4	5	6+
0	.175	.158	.227	.889	0	0
	463	480	22	9	8	9
	.851	.980	1.066	2.667	0	0
1	3.594	4.674	5.498	6.531	6.289	7.160
	537	771	271	224	90	81
	3.357	3.973	3.866	5.232	5.106	4.641
2	4.533	6.437	8.887	9.606	10.871	10.15
	15	453	256	236	170	120
	2.973	4.475	5.700	6.149	7.104	7.140
3		10.90	10.0	11.750	14.913	16.029
		10	43	68	23	34
		5.801	6.962	7.795	9.709	11.221
4+		6.250	10.636	13.097	20.918	18.381
		4	11	31	11	21
		6.131	6.682	7.871	9.704	9.836

The following information is reported in each cell:

Average household trip rate $\bar{t}_{ij}$

Number of sampled households in cell n_{ij}

Std. Deviation of trip rates in cell s_{ij}

had less than 30 observations (only exception is cell 4,4, with $n_{4,4} = 31$). It was therefore decided to combine these cells with the corresponding cells with "3" cars. No comparable decision could be made with respect to the household size independent variable since most of the "6⁺" household size cells had enough observations to warrant more refined statistical analysis.

The cross-classification table was therefore reduced to a table with 0, 1, 2, 3⁺ cars and 1, 2, 3, 4, 5, 6⁺ household size. It is reproduced in Table 34.

Step 3: Pairwise comparisons of the means of the following cells were made: 1,5 vs. 1,6; 2,5 vs. 2,6; 3,5 vs. 3,6. Cells with zero cars were not tested due to the insignificant number of samples in these cells, which makes their rates unreliable (these cells are also relatively unimportant in terms of their contribution to overall vehicle traffic considering the low magnitude of their rates). These comparisons were used to decide whether the "6⁺" level of household size resulted in rates significantly higher than those in the corresponding "5" size level cells. As previously mentioned, t-tests were used for this purpose. The test statistic appropriate for the case at hand (comparing two means for cells (i,j) and (i,j+1), containing a sample from independent random variables $t_{i,j}$ and $t_{i,j+1}$ respectively, where the respective sample sizes $n_{i,j}$ and $n_{i,j+1}$ are unequal and greater than 30, and the respective population variances $\sigma_{i,j}^2$ and $\sigma_{i,j+1}^2$ are unknown and not necessarily equal) is (47):

$$t^* = \frac{\bar{t}_{i,j+1} - \bar{t}_{i,j}}{\sqrt{(S_{i,j+1}^2/n_{i,j+1}) + (S_{i,j}^2/n_{i,j})}}$$

Table 34. Cross-Classification of Total Vehicle Trip Rates, LAFAYETTE
4 Levels of Auto Ownership and 6 Levels of Household Size

Autos Per Household	Household Size					
	1	2	3	4	5	6+
0	.175	.158	.227	.889	0	0
	463	480	22	9	8	9
	.851	.980	1.066	2.667	0	0
1	3.594	4.674	5.498	6.531	6.289	7.160
	537	771	271	224	90	81
	3.357	3.973	3.866	5.232	5.106	4.641
2	4.533	6.437	8.887	9.606	10.871	10.15
	15	453	256	236	170	120
	2.973	4.475	5.700	6.149	7.104	7.140
3+		9.571	10.130	12.172	16.824	16.927
		14	54	99	34	55
		6.060	6.849	7.804	9.962	10.682

The following information is reported in each cell:

Average household trip rate $\bar{t}_{ij}$

Number of sampled households in cell n_{ij}

Std. Deviation of trip rates in cell s_{ij}

The hypothesis H_0 is: "The two means are equal", and the alternative hypothesis H_1 is that the mean corresponding to a higher level of the independent variable is greater. The test statistic t^* is compared to the value of the t-distribution with the appropriate degrees of freedom (d.f.) and significance level (α). The degrees of freedom are computed using the following expression:

$$d.f. = \frac{[(S_{i,j+1}^2/n_{i,j+1}) + (S_{i,j}^2/n_{i,j})]^2}{\frac{(S_{i,j+1}^2/n_{i,j+1})^2}{n_{i,j+1} + 1} + \frac{(S_{i,j}^2/n_{i,j})^2}{n_{i,j} + 1}} - 2$$

The significance level was set at 0.05. The hypothesis H_0 that the means were equal would be rejected only when $t^* \geq t_{0.05,d.f.}$. Standard statistical tables are used to determine $t_{0.05,d.f.}$.

A sample test calculation is shown in Table 35. A summary of the t-tests for the three pairs of means is shown in Table 36. None of the three tests revealed any significant differences between the paired means. Therefore households with 6 or more members do not produce more total vehicle trips than those with 5 members and the same number of cars. The classification frame was therefore reduced to a table with 0, 1, 2, 3⁺ cars and 1, 2, 3, 4, 5⁺ household size. It is reproduced in Table 37.

Step 4: More comparisons were made to test if households with 5 or more members produce a higher number of vehicle trips than those households with 4 members and the same auto ownership level. The same tests as in the preceding step were run, comparing the means of the following cells: 1,4 vs. 1,5; 2,4 vs. 2,5; 3,4 vs. 3,5. As in the previous step, "0"-cars cells were not tested due to the insignificant number of samples in

Table 35. Sample Test Calculation - Comparison of the Means of Cells (CARS = 1, SIZE = 5) and (CARS = 1, SIZE = 5)

$$H_0 : \mu_{1,6} = \mu_{1,5}$$

$$H_1 : \mu_{1,6} > \mu_{1,5}$$

$$\alpha = 0.05$$

Decision Rule: $t^* \geq t_{\alpha, d.f.}$

Reject H_0

$$\text{where } t^* = \frac{\bar{t}_{1,6} - \bar{t}_{1,5}}{\sqrt{\frac{s_{1,6}^2}{n_{1,6}} + \frac{s_{1,5}^2}{n_{1,5}}}}$$

$$\text{and d.f.} = \frac{[(s_{1,6}^2/n_{1,6}) + (s_{1,5}^2/n_{1,5})]^2}{\frac{(s_{1,6}^2/n_{1,6})^2}{n_{1,6} + 1} + \frac{(s_{1,5}^2/n_{1,5})^2}{n_{1,5} + 1}} - 2$$

$$\bar{t}_{1,6} = 7.160$$

$$s_{1,6} = 4.641$$

$$n_{1,6} = 81$$

$$\bar{t}_{1,5} = 6.289$$

$$s_{1,5} = 5.106$$

$$n_{1,5} = 90$$

$$t^* = \frac{7.160 - 6.289}{\sqrt{\frac{(4.641)^2}{81} + \frac{(5.106)^2}{90}}} = 1.168$$

$$d.f. = \frac{[\frac{(4.641)^2}{81} + \frac{(5.106)^2}{90}]^2}{\frac{[(4.641)^2/81]^2}{82} + \frac{[(5.106)^2/90]^2}{91}} - 2 = 170$$

$$t_{0.05, 170} = 1.655 < t^*$$

Accept H_0

Table 36. Summary of t-tests for Step 3

$$H_0: \mu_{i,j} = \mu_{i,j+1} \quad H_1: \mu_{i,j+1} > \mu_{i,j} \quad \alpha = 0.05$$

<u>Cell (i,j)</u>	vs.	<u>Cell (i,j+1)</u>	<u>t*</u>	<u>d.f.</u>	<u>t_{0.05,d.f.}</u>	<u>Decision</u>
1,5		1,6	1.168	170	1.655	Accept H ₀
2,5		2,6	-0.845	257	1.653	Accept H ₀
3,5		3,6	0.046	76	1.66	Accept H ₀

None of the pairwise differences are significant.

Table 37. Cross-Classification of Total Vehicle Trip Rates, LAFAYETTE:
4 Levels of Auto Ownership and 5 Levels of Household Size

Autos Per Household	Household Size				
	1	2	3	4	5+
0	.175	.158	.227	.889	0
	463	480	22	9	17
	.851	.980	1.066	2.667	0
1	3.594	4.674	5.498	6.531	6.702
	537	771	271	224	171
	3.357	3.973	3.866	5.232	4.896
2	4.533	6.437	8.887	9.606	10.572
	15	453	256	236	290
	2.973	4.475	5.70	6.149	7.116
3+		9.571	10.130	12.172	16.888
		14	54	99	89
		6.060	6.849	7.804	10.356

The following information is reported in each cell:

Average household trip rate $\bar{t}_{ij}$

Number of sampled households in cell n_{ij}

Std. deviation of trip rates within cell s_{ij}

these cells. The above tests are summarized in Table 38. In two of the three cases, the hypothesis was rejected with 95 percent confidence. It can be concluded that households of 5 or more members and owning one car only do not produce significantly more total vehicle trips than households of 4 which own one car only. However, this is not the case for auto ownership levels 2 and 3, at which households of 5 or more members do make significantly more trips than those of 4 members. Since in two out of three cases there was no statistical evidence to warrant combining the "4" and "5⁺" household size levels, it was decided that "5⁺" will be the highest level of that variable.

Step 5: The needed level of disaggregation of the auto ownership variable was investigated by comparing the (2,j) cell-means with the corresponding (3,j) cell means. In other words, the hypothesis is that households of certain size and owning 3 or more cars do not produce more total vehicle trips than households of the same size but with only 2 cars. The following pairs of cells (with significant numbers of observations) were compared: 2,3 vs. 3,3; 2,4 vs. 3,4; 2,5 vs. 3,5. These tests are summarized in Table 39. In two of the three cases, the hypothesis was rejected with 95 percent confidence. It can be concluded

Table 38. Summary of t-tests for Step 4

$$H_0: \mu_{i,j} = \mu_{i,j+1} \quad H_1: \mu_{i,j+1} > \mu_{i,j} \quad \alpha = 0.05$$

Cell(i,j)	vs.	Cell(i,j+1)	t*	d.f.	t _{0.05,d.f.}	Decision
1,4		1,5	0.334	379	1.65	Accept H ₀
2,4		2,5	1.669	524	1.646	Reject H ₀
3,4		3,5	3.495	164	1.65	Reject H ₀

Only one of the differences (1,4 vs. 1,5) is not significant.

Table 39. Summary of t-tests for Step 5

$$H_0: \mu_{2,j} = \mu_{3,j} \quad H_1: \mu_{3,j} > \mu_{2,j} \quad \alpha = 0.05$$

Cell(2,j)	vs.	Cell(3,j)	t*	d.f.	$t_{0.05,d.f.}$	Decision
2,3		3,3	1.246	70	1.668	Accept H_0
2,4		3,4	2.91	152	1.66	Reject H_0
2,5		3,5	5.36	115	1.66	Reject H_0

Only one of the differences (2,3 vs. 3,3) is not significant.

that households of size 4, 5 or more, and owning three or more cars produce significantly more total vehicle trips than those same sized households that own only two cars. For households of size 3, it seems that owning 3 or more cars does not lead to a significant increase in overall trip making as compared to owning only 2 cars. The latter hypothesis could have been rejected with an α of 0.2, that is with 80 percent confidence, since $t_{0.2,70} = 0.847$ which is less than the value of the t-statistic calculated for that comparison (1.246). Therefore, since in two of the three cases, the hypothesis of equality of all means has been rejected with 95 percent confidence, and in the third case, it could be rejected with 80 percent confidence, it was determined that the "3⁺" will be the highest level of the auto ownership variable.

The final conclusion is that the following level of disaggregation is needed for the stability of trip rates:

CARS: 0, 1, 2, 3⁺

HOUSEHOLD SIZE: 1, 2, 3, 4, 5⁺

It should be noted that in the above statistical analysis, all tests were made on a cell-by-cell basis rather than on level-by-level (of the independent variable). The reason for this longer procedure is that it was felt that testing the overall mean for a level of a certain independent variable would mask the special interactions between the independent variables that can only be detected at the individual cell level.

No further testing was made as to whether other rows or columns could be combined. However, the same test presented above can be used for any pairwise comparison of cell means. The reason for investigating only the higher levels of the independent variables is that it is generally accepted that trip rates will vary for these different levels of

the independent variables, and that it is only after a certain level that the trip rates will tend to stabilize. Moreover, for these cells corresponding to the more frequent households, the average trip rate is usually fairly reliable, and there is no reason for not using the data available to its fullest by refining it as much as is practically and efficiently feasible.

The number of levels of each of the independent variables determined in this section will be used for the development of cross-classified household vehicle trip rates for two of the study areas as described in the next section.

Cross-Classified Vehicle Trip Rate Tables for Lafayette and Evansville

As discussed in Chapter 2, appropriate disaggregate household data was available only in two of the study areas, Lafayette and Evansville. Household cross-classified trip rates for each purpose (HBW,HBO,NHB) were developed for each of these urban areas. The following had to be done in both cases:

1. Processing of the data files containing the raw O-D data. Trips for each purpose were accumulated for all the surveyed households.
2. Trips were accumulated within each household socio-economic category, and average trip rates, as well as the standard deviations were computed for each category.

The trip rate tables obtained are shown in Tables 40 to 42 for Lafayette (HBW,HBO and NHB respectively), and in Tables 43 to 46 for Evansville (HBW, HBO, NHB and ALL respectively). All-purposes trip rates for Lafayette have already been shown in Table 37. Plots of the trip rates for the two areas for all four types of trips are shown in

Table 40. Cross-Classification of HBW Vehicle Trip Rates, LAFAYETTE

<u>Autos Per Household</u>	<u>Household Size</u>				
	1	2	3	4	5+
0	0.004	0.004	0.091	0	0
	463	480	22	9	17
	0.093	0.064	0.426	0	0
1	0.572	0.868	1.118	1.371	1.456
	537	771	271	224	171
	1.129	1.30	1.339	1.507	1.390
2	0.867	1.870	2.266	2.280	2.231
	15	453	256	236	290
	1.457	1.809	1.815	1.642	1.667
3+		2.00	2.704	2.232	2.888
		14	54	99	89
		2.075	2.376	2.272	2.323

The following information is reported in each cell:

Average household trip rate $\bar{t}_{ij}$

Number of sampled households in cell n_{ij}

Std. Deviation of trip rates in cell s_{ij}

Table 41. Cross-Classification of HBO Vehicle Trip Rates, LAFAYETTE

Autos Per Household	Household Size				
	1	2	3	4	5+
0	0.112	0.092	0.091	0.444	0
	463	480	22	9	17
	0.624	0.630	0.426	1.333	0
1	2.019	2.459	2.756	3.192	3.333
	537	771	271	224	171
	2.157	2.438	2.537	3.103	3.157
2	2.933	2.656	3.746	4.119	5.083
	15	453	256	236	290
	2.251	2.524	3.229	3.263	4.003
3		3.714	4.685	7.071	7.993
		14	54	99	89
		2.813	4.013	5.496	5.006

The following information is reported in each cell:

Average household trip rate $\bar{t}_{ij}$

Number of sampled households in cell n_{ij}

Std. Deviation of trip rates in cell s_{ij}

Table 42. Cross-Classification of NHB* Vehicle Trip Rates, LAFAYETTE

Autos Per Household	Household Size				
	1	2	3	4	5+
0	0.058	0.063	0.045	0.444	0
	463	480	22	9	17
	0.430	0.537	0.218	1.333	0
1	1.004	1.348	1.624	1.969	1.912
	537	771	271	224	171
	1.733	2.194	2.124	2.886	2.652
2	0.733	1.912	2.875	3.208	3.259
	15	453	256	236	290
	1.280	2.582	3.483	3.772	3.608
3+		3.857	2.741	2.869	6.067
		14	54	99	89
		4.240	3.343	3.132	6.239

The following information is reported in each cell:

Average household trip rate $\bar{t}_{ij}$

Number of sampled households in cell n_{ij}

Std. Deviation of trip rates in cell s_{ij}

*Truck trips are not included.

Table 43. Cross-Classification of HBW Vehicle Trip Rates, EVANSVILLE

Autos Per Household	Household Size				
	1	2	3	4	5+
0	0.006	0.008	0	0.034	0
	482	250	56	29	45
	0.102	0.126	0	0.186	0
1	0.829	0.910	1.179	1.375	1.345
	381	934	329	299	287
	1.140	1.109	1.143	1.150	1.111
2	1.00	1.982	1.894	2.044	2.023
	11	610	433	476	488
	0.894	1.408	1.270	1.339	1.329
3+		1.929	2.812	2.967	2.973
		42	117	92	146
		1.688	1.833	1.788	2.071

The following information is reported in each cell:

Average household trip rate $\bar{t}_{ij}$

Number of sampled households in cell n_{ij}

Std. Deviation of trip rates in cell s_{ij}

Table 44. Cross-Classification of HBO Vehicle Trip Rates, EVANSVILLE

Autos Per Household	Household Size				
	1	2	3	4	5+
0	0.006	0.032	0	0.138	0
	482	250	56	29	45
	0.137	0.309	0	0.743	0
1	1.430	1.941	2.623	2.696	3.376
	381	934	329	299	287
	1.784	2.093	2.459	2.591	3.368
2	2.00	2.315	3.464	3.868	4.553
	11	610	433	476	488
	1.673	2.171	2.595	2.916	3.466
3+		2.381	4.265	5.717	5.815
		42	117	92	146
		2.118	3.033	3.929	3.692

The following information is reported in each cell:

Average household trip rate $\bar{t}_{ij}$

Number of sampled households in cell n_{ij}

Std. Deviation of trip rates in cell s_{ij}

Table 45. Cross-Classification of NHB* Vehicle Trip Rates, EVANSVILLE

Autos Per Household	Household Size				
	1	2	3	4	5+
0	0	0.024	0	0	0
	482	250	56	29	45
	0	0.282	0	0	0
1	0.950	1.005	1.532	1.441	1.481
	381	934	329	299	287
	1.630	1.735	2.367	1.990	2.070
2	1.182	1.966	2.527	2.653	2.670
	11	610	433	476	488
	1.250	3.117	3.128	3.267	3.504
3+		2.690	3.359	3.565	3.747
		42	117	92	146
		2.867	3.883	5.366	4.017

The following information is reported in each cell:

Average household trip rate $\bar{t}_{ij}$

Number of sampled households in cell n_{ij}

Std. Deviation of trip rates in cell s_{ij}

*Truck trips are not included

Table 46. Cross-Classification of ALL* Vehicle Trip Rates, EVANSVILLE

Autos Per Household	Household Size				
	1	2	3	4	5+
0	0.012	0.064	0	0.172	0
	482	250	56	29	45
	0.204	0.563	0	0.759	0
1	3.210	3.857	5.334	5.512	6.202
	381	934	329	299	287
	3.066	3.197	4.047	3.778	4.430
2	4.182	6.262	7.885	8.565	9.246
	11	610	433	476	488
	2.228	4.385	4.698	5.074	5.601
3+		7.00	10.436	12.250	12.534
		42	117	92	146
		3.939	5.520	7.249	6.296

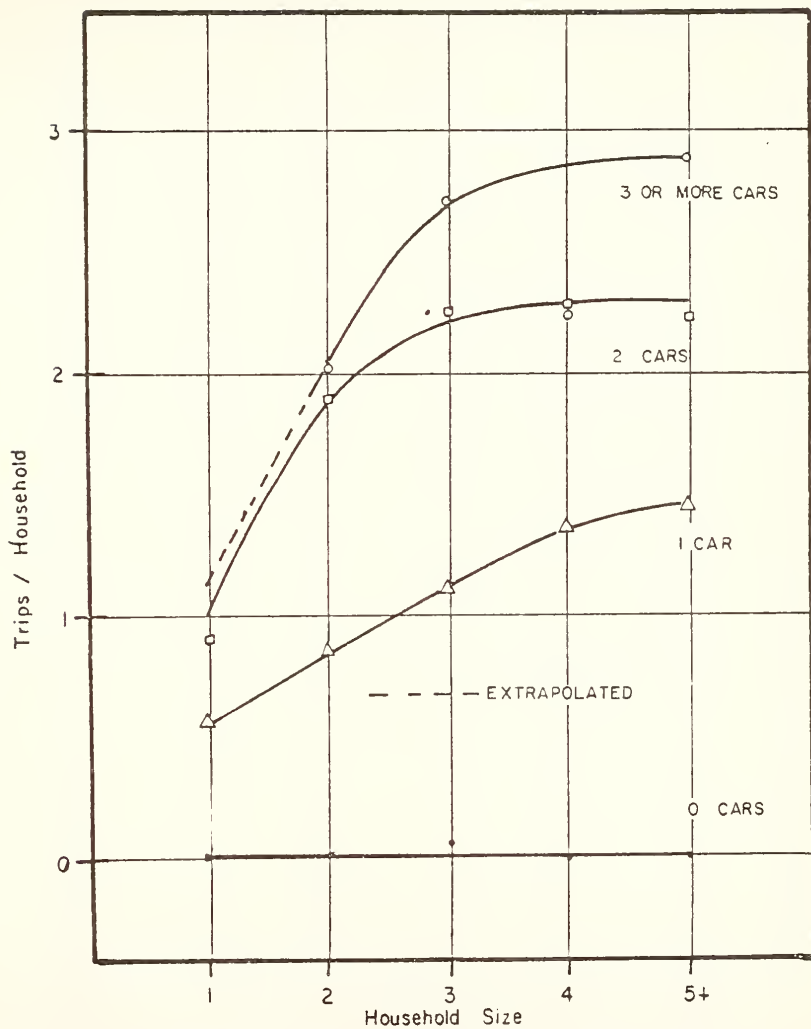
The following information is reported in each cell:

Average household trip rate $\bar{t}_{ij}$

Number of sampled households in cell n_{ij}

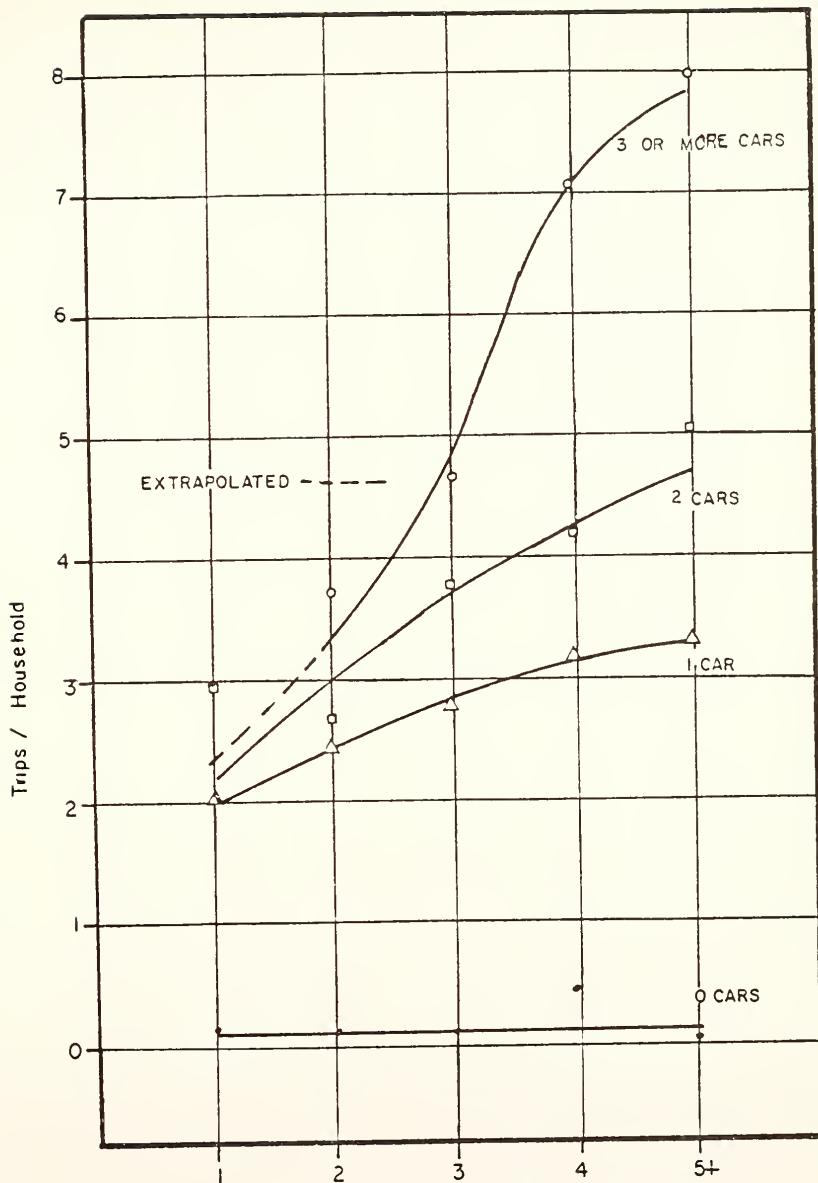
Std. Deviation of trip rates in cell s_{ij}

*Truck trips are not included



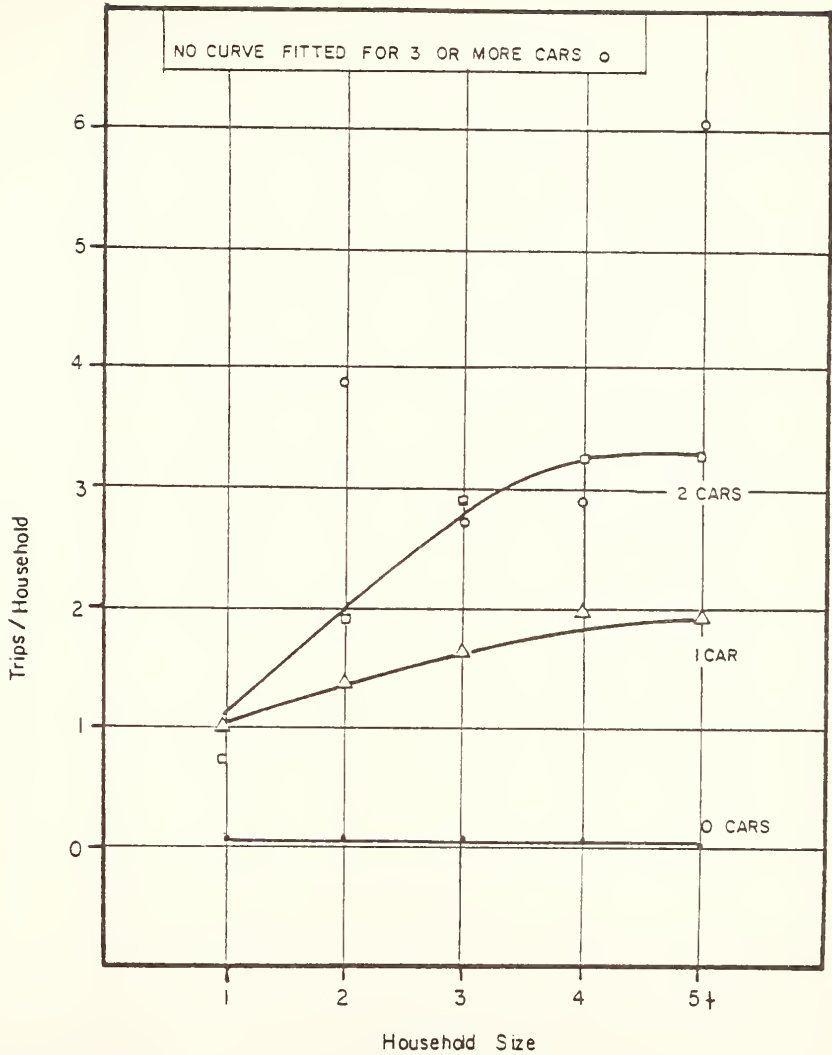
HOME-BASED WORK VEHICLE TRIPS PER HOUSEHOLD BY
HOUSEHOLD SIZE AND AUTO OWNERSHIP-LAFAYETTE

FIGURE 3



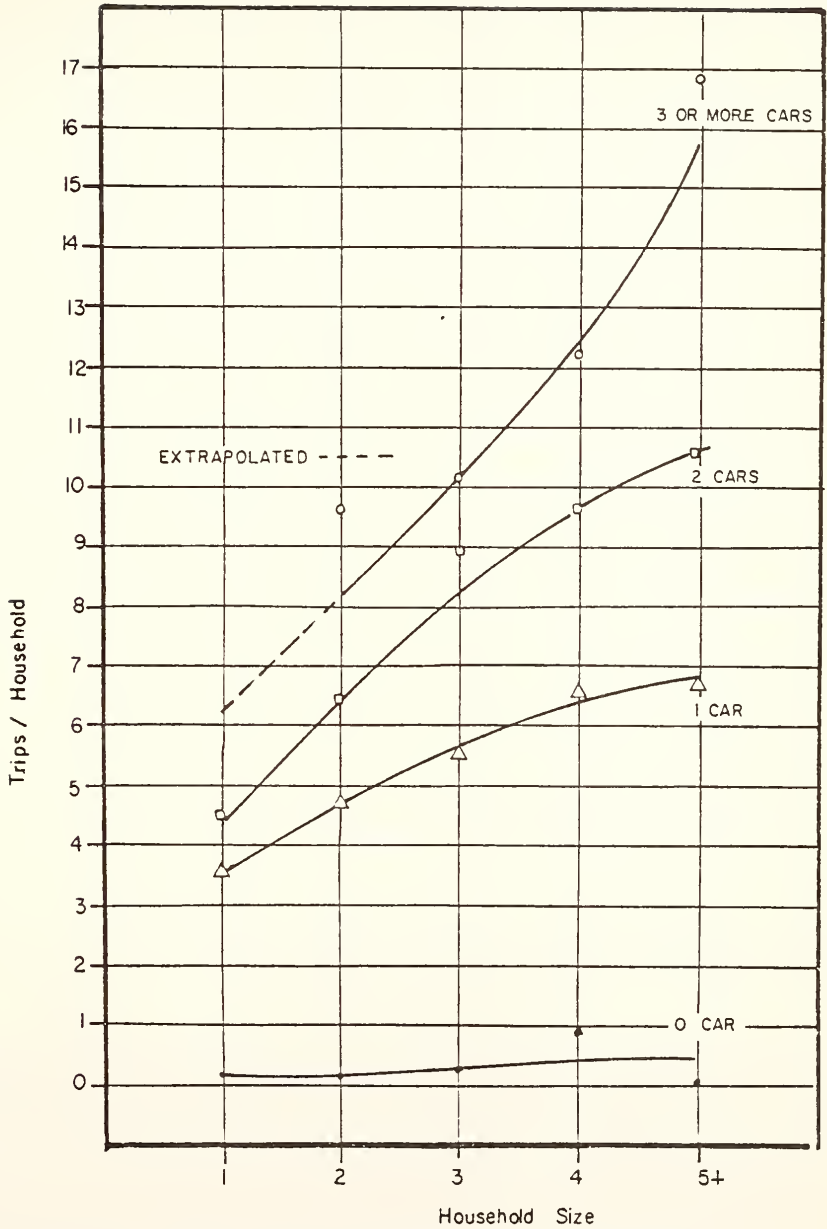
HOME-BASED OTHER VEHICLE ^{Household Size} TRIPS PER HOUSEHOLD
BY HOUSEHOLD SIZE AND AUTO OWNERSHIP—LAFAYETTE

FIGURE 4



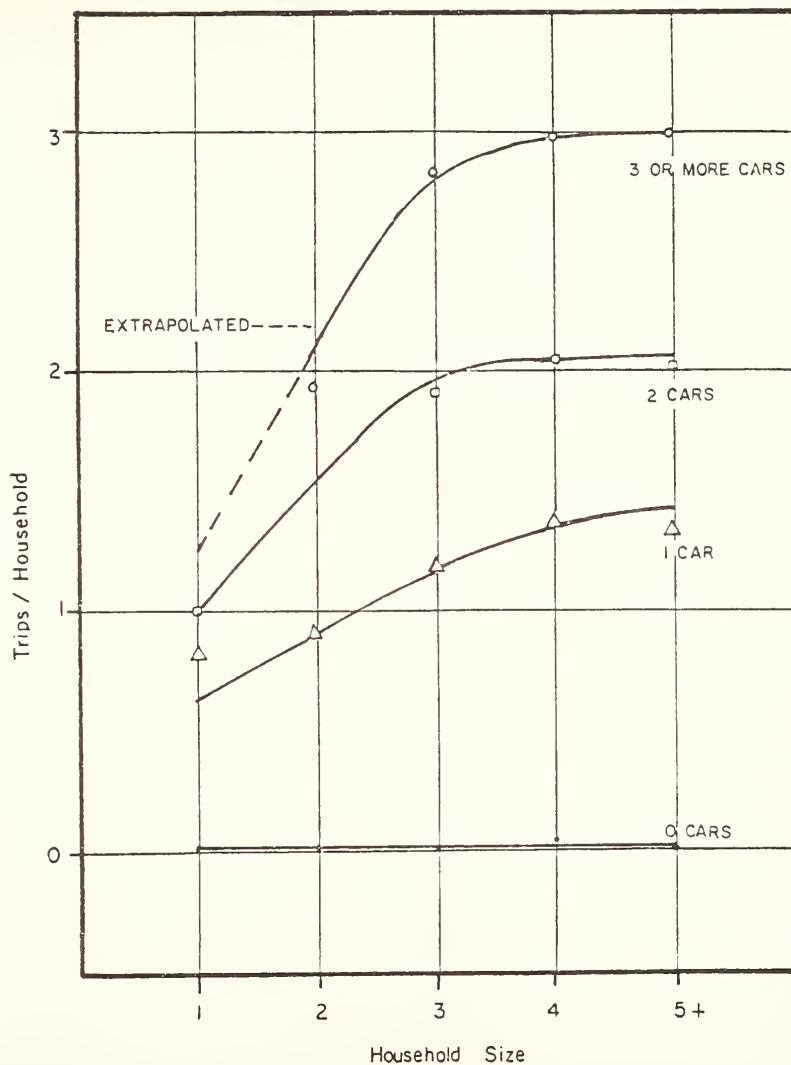
NON-HOME BASED VEHICLE TRIPS PER HOUSEHOLD BY
HOUSEHOLD SIZE AND AUTO OWNERSHIP—LAFAYETTE

FIGURE 5



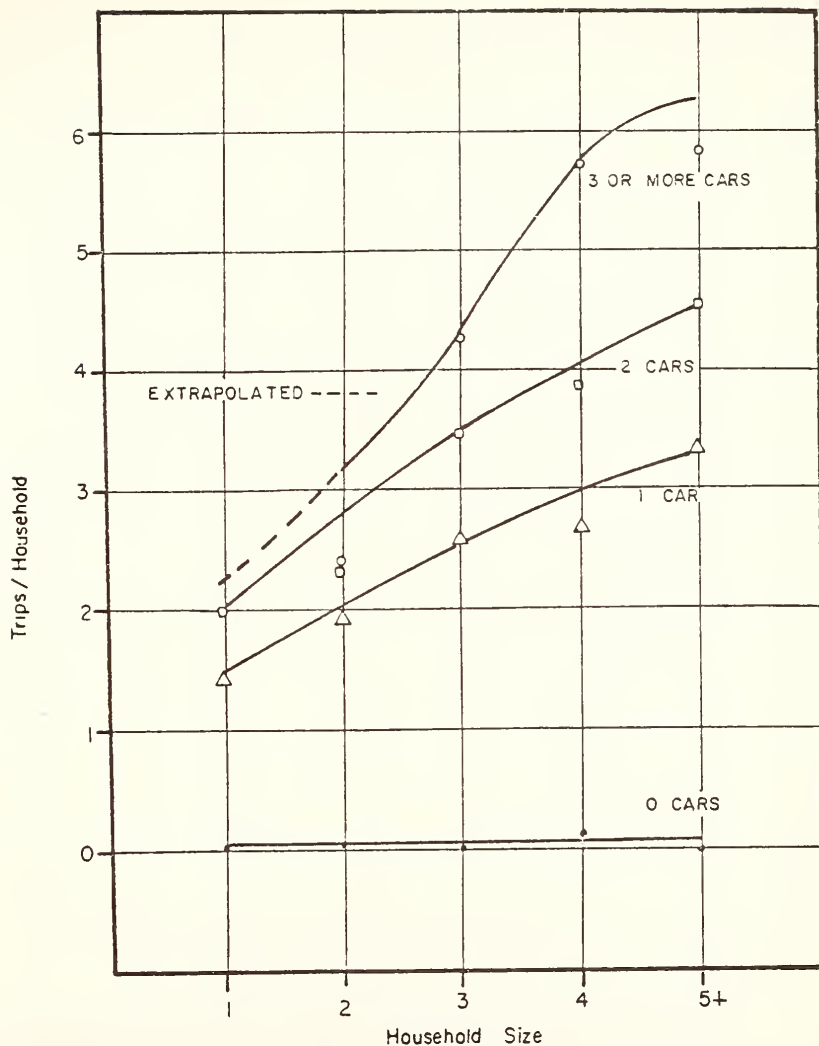
ALL VEHICLE TRIPS PER HOUSEHOLD BY HOUSEHOLD
SIZE AND AUTO OWNERSHIP - LAFAYETTE

FIGURE 6



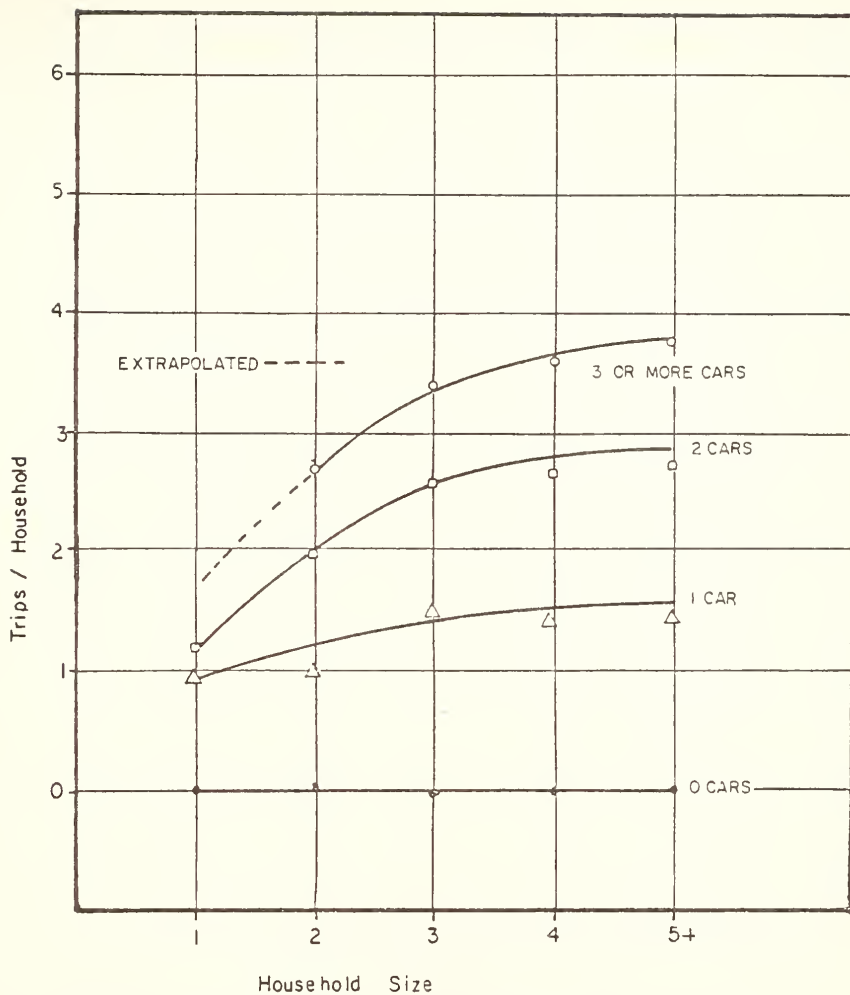
HOME-BASED WORK VEHICLE TRIPS PER HOUSEHOLD BY
HOUSEHOLD SIZE AND AUTO OWNERSHIP - EVANSVILLE

FIGURE 7



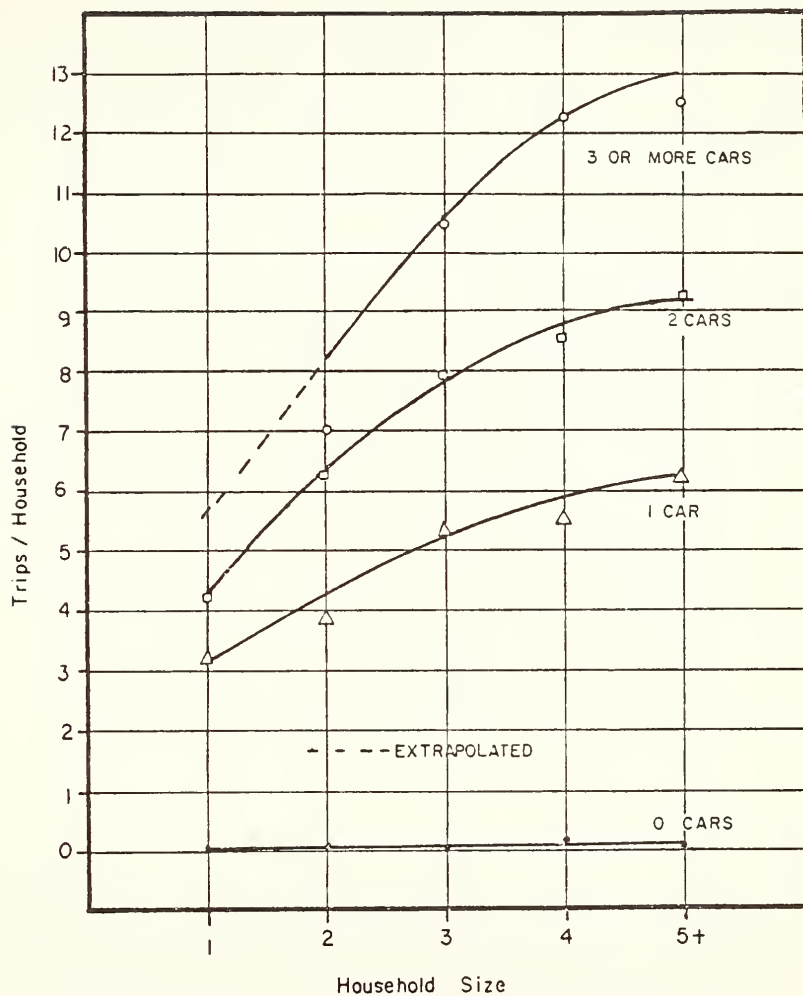
HOME-BASED OTHER VEHICLE TRIPS PER HOUSEHOLD BY
HOUSEHOLD SIZE AND AUTO OWNERSHIP—EVANSVILLE

FIGURE 8



NON-HOME BASED VEHICLE TRIPS PER HOUSEHOLD BY
HOUSEHOLD SIZE AND AUTO OWNERSHIP - EVANSVILLE

FIGURE 9



ALL VEHICLE TRIPS PER HOUSEHOLD BY HOUSEHOLD
SIZE AND AUTO OWNERSHIP - EVANSVILLE

FIGURE 10

Figures 3 to 6 for Lafayette and 7 to 10 for Evansville. It should be noted that for prediction purposes, it is advisable to read the appropriate trip rates off the graphs rather than directly from the tables. Indeed, the graphs have tried to correct for bad data points and to ensure that the relationship developed in the cross-classification is logical by fitting a smooth curve through the observations.

The cross-classified trip rate tables developed for the two study areas will be used to evaluate the transferability of household trip generation models between small urban areas in view of their use in synthetic modeling efforts. Before doing so, a brief discussion of the temporal and spatial stability of these models will be presented.

Temporal and Spatial Stability of Household Trip Generation Models

It was mentioned earlier in this chapter that temporal and spatial stability of disaggregate household trip generation models were "controversial" advantages of these models over aggregate zonal regression models.

Concerning temporal stability, earlier studies had made a strong case for the stability of household trip generation models over time, thus forming a stronger basis for future predictions (44). However, other studies did cast some doubts on the matter, even though no definite conclusions were made as to their non-stability (48). In the final analysis, limited tests which are specific to one particular geographic area cannot provide a universal generalized conclusion concerning that issue. The fact remains however that the issue of temporal stability is one that is associated with all predictive models, and disaggregate household models, because of their behavioral nature, remain superior in that respect to zonal aggregate regression models. More importantly,

cross-classification models offer the advantage of being easy to monitor and update.

As far as spatial stability is concerned, the same controversy exists. However, this property is crucial to synthetic modeling applications which advocate borrowing travel relationships from other urban areas. Overall, results have been encouraging. Experiments in England have shown positive results as to the ability of category models developed in a certain city to synthesize travel patterns in another city (25). Studies in the Southeastern Wisconsin Region have shown that the difference in trip rates between large and small urban areas was not significant except for the non-home based trip rates of Milwaukee and Kenosha, and the home based shopping trip rates of Milwaukee and Racine (24). These comparisons were made using a test statistic which was compared to the Chi-Square distribution. That test statistic (Q-statistic) was developed specifically for that purpose, and the details of it can be found in references 16, 24 or 49. That same testing technique was used to compare person trip rates from Elizabethton, Tennessee; Murray, Kentucky; Paducah, Kentucky and Racine, Wisconsin (all with populations between 20,000 and 50,000) (16). The results of these tests indicated that in most cases the trip rates from different cities were significantly different.

The above conclusions however do not imply that trip rates are not transferable between similar urban areas. For one, the statistical testing procedure used in that study (Q-Statistic) might not be all that appropriate for the comparison it was applied to. That test provides an overall comparison of all trip rates from two separate tables.

It is however very sensitive to differences of individual cell mean pairs. In other words, if the rates for one of the cells are somewhat different for the two tables, the test might lead to the conclusion that the two tables are significantly different. Since only some cells might have different average trip rates (many times due to bad data points or slight deviations in each of the cells), trip tables should be compared on a cell-by-cell basis. Cell-by-cell comparisons of the cross-classified trip rates developed for each purpose for Lafayette and Evansville are described in the next section.

Comparison of Cross-Classified Household Trip Rates for Lafayette and Evansville

The cell-by-cell comparisons can be made using a series of t-tests. The test statistic used in this situation involves the comparison of two means for two cells each containing a sample from two independent random variables, where the respective sample sizes are unequal and greater than 30, and the respective population variances are unknown and not necessarily equal. The significance of the difference between the cell means from the two urban areas is tested. The null hypothesis H_0 is that the difference is not significant, while the alternative hypothesis H_1 is that this difference is significant. The computational formulas are shown in Table 47.

The results of these tests are summarized in Tables 48, 49 and 50 for Home Based Work, Home Based Other and Non-Home Based vehicle trips respectively. Only a few cells for all three purposes exhibited significant differences between the two urban areas. Referring to Figures 3 to 11, it can be observed that the cells that exhibit major differences

Table 47. T-test for Comparison of Cell Means

X and Y Independent random variables: trips per household in same category for Lafayette and Evansville, respectively.

No distribution assumed.

n_x and n_y Respective sample sizes in that category.

σ_x and σ_y Unknown and not necessarily equal

$$H_0 : \mu_x = \mu_y \quad H_1 : \mu_x \neq \mu_y \quad \alpha = 0.05$$

Decision Rule: $|t^*| \geq t_{\alpha/2, d.f.}$ Reject H_0

where
$$t^* = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$$

(s_x^2 and s_y^2 are respective sample variances)

and
$$d.f. = \frac{[(s_x^2/n_x) + (s_y^2/n_y)]^2}{\frac{(s_x^2/n_x)^2}{n_x + 1} + \frac{(s_y^2/n_y)^2}{n_y + 1}} - 2$$

Table 48. Cell-by-Cell Comparison of Lafayette & Evansville

HBW Trip Rate Tables: Summary of t-tests Results

CELLS		d.f.	t*	$t_{d.f.,0.025}$	Are means significantly different?
CARS	SIZE				
0	1	942	0.3152	1.960	No
0	2	318	0.4713	1.970	No
1	1	815	3.379	1.960	Yes
1	2	1522	0.709	1.960	No
1	3	535	0.5929	1.960	No
1	4	404	0.0331	1.967	No
1	5	300	0.8887	1.970	No
2	2	828	1.0944	1.960	No
2	3	405	2.8879	1.967	Yes
2	4	394	1.9060	1.967	No
2	5	507	1.8103	1.965	No
3	3	843	0.2959	1.960	No
3	4	186	2.4935	1.974	Yes
3	5	171	0.2833	1.974	No

Table 49. Cell-by-Cell Comparison of Lafayette and Evansville
HBO Trip Rate Tables: Summary of t-tests Results

CELLS		d.f.	t*	$t_{d.f.,0.025}$	Are means significantly different?
CARS	SIZE				
0	1	505	3.5734	1.965	Yes
0	2	728	1.7257	1.960	No
1	1	897	4.5151	1.960	Yes
1	2	1528	4.6519	1.960	Yes
1	3	571	0.648	1.960	No
1	4	431	1.939	1.967	No
1	5	378	0.1375	1.968	No
2	2	888	2.3101	1.960	Yes
2	3	450	1.1887	1.965	No
2	4	426	1.0002	1.967	No
2	5	542	1.8753	1.960	No
3	3	83	0.6842	1.99	No
3	4	179	1.9689	1.974	No
3	5	147	3.459	1.978	Yes

Table 50. Cell-by-Cell Comparison of Lafayette & Evansville
NHB Trip Rate Tables: Summary of t-tests Results

<u>CELLS</u>		<u>d.f.</u>	<u>t*</u>	<u>t_{d.f.,0.025}</u>	<u>Are Means Significantly Different?</u>
<u>CARS</u>	<u>SIZE</u>				
0	1	462	2.9024	1.965	Yes
0	2	730	1.2866	1.960	No
1	1	850	0.4817	1.960	No
1	2	1453	3.5254	1.960	Yes
1	3	596	0.5013	1.960	No
1	4	376	2.3513	1.968	Yes
1	5	294	1.8203	1.970	No
2	2	1050	0.3085	1.960	No
2	3	492	1.3155	1.965	No
2	4	415	1.9298	1.967	No
2	5	595	2.2254	1.960	Yes
3	3	121	1.0664	1.981	No
3	4	145	1.0842	1.979	No
3	5	134	3.1343	1.980	Yes

are ones whose rates do not fall on the fitted curves. Therefore, the significant differences indicated in some cases can be attributed to the inaccuracy of some of the observations which might have led to an erroneous rate.

The above comparison strengthens the belief that behavioral disaggregate household models are transferable from one urban area to another, especially in the case of small urban areas, regardless of the socioeconomic differences between them. Indeed, previous comparisons in this study have shown Lafayette and Evansville to be different in the distribution of households between auto ownership categories (c.f. Chapter 4), in the areawide trip purpose percentages (c.f. Chapter 3), as well as in the areawide trip frequency rates (c.f. Chapter 4). Moreover, zonal regression trip generation equations developed in Lafayette did not perform well in Evansville (c.f. Chapter 5). In terms of population, even though Evansville and Lafayette are in the same population size group, Evansville is considerably larger. In spite of all these differences, the household trip rates were not significantly different, which means that the trip making behavior is in essence similar in these cities even though their aggregate travel patterns are different.

The above conclusions clearly indicate the superiority of household category models as tools for synthetic trip generation analysis within the conventional urban transportation planning structure.

CHAPTER 7: BAYESIAN UPDATING IN TRANSPORTATION PLANNING

The previous chapters have investigated various synthetic modeling techniques that can be used for trip generation analysis in small urban areas. These techniques consisted mainly of borrowing models developed in similar areas, such as borrowing cross-classified household trip rates. Even though these trip rates were shown to be stable between small urban areas, it might be desirable to modify them so that they reflect local conditions more accurately. This can be done by updating the borrowed parameters using information particular to the area under study. This information could be objective or subjective. Objective information would typically be derived from a small sample collected on site, which is a relatively modest effort compared to a conventional O-D survey. Subjective information is that which reflects the feeling of and experience accumulated by experts familiar with the study area. Some updating procedures will be briefly reviewed in this chapter, with special emphasis being placed on one of them, Bayesian updating. Previous applications of Bayesian statistics in the transportation field will be also discussed, and a methodology for their use to update cross-classified trip rates will be developed. This methodology will consider the following cases:

- A small disaggregate sample is collected in the area under study
- Consecutive small samples are collected as part of a continuing planning process

- Subjective information is used to update the trip rates.

Detailed step by step applications will be presented for each of the above cases.

Updating Procedures

Alternative approaches for updating disaggregate travel models have been described and evaluated by Ben Akiva and Atherton (50). These can be summarized as follows:

1. Model Constants (as opposed to model coefficients of the independent variables) are adjusted to better reflect the new study area using aggregate data from that area.
2. All model parameters are re-estimated using a small (disaggregate) sample collected in the study area.
3. Only the constant is re-estimated using the small sample.
4. A methodology based on Bayesian statistics is used to update the original model using information from the study area. The test showed this technique to be superior to the others.

Use of Bayesian Updating in Transportation

Bayesian updating is used when prior information (such as a model from a certain city) is to be updated using new information (such as a small sample). The updated information is called the posterior information. In probabilistic terminology, Bayes' theorem can be formulated as:

$$\left[\begin{array}{l} \text{Posterior Probability} \\ \text{of parameter } \theta, \\ \text{given the sample} \\ \text{outcome} \end{array} \right] = \left[\begin{array}{l} \text{normalizing} \\ \text{constant} \end{array} \right] \times \left[\begin{array}{l} \text{Sample} \\ \text{Likelihood,} \\ \text{given } \theta \end{array} \right] \times \left[\begin{array}{l} \text{Prior} \\ \text{probability} \\ \text{of } \theta \end{array} \right]$$

Formulas and expressions for posterior information parameters have been derived from the basic theorem for application to various general situations (51). Such situations have been identified in some transportation-related areas, such as an effort by Isibor (1969) to adapt Bayesian updating techniques in the context of modeling the impact of highway improvements on the value of adjacent land parcels (52).

In the transportation planning context, Sinha (53) showed how Bayesian statistics could be used by the planner to arrive at an optimal decision by combining subjective evaluation with objective information. He presented an actual application of that technique to the problem of county-level population forecasting in the southeastern Wisconsin region.

Bayesian statistics were also used to update objective information using subjective judgement for the purpose of mode choice estimation in small urban areas (54). Chan et al. (18) applied Bayesian methodology to update travel demand elasticities. An application has also been demonstrated in the field of pavement management (55). These applications demonstrate the flexibility and the tremendous potential of Bayesian updating techniques.

Bayesian Updating of Cross-Classified Trip Generation Rates

Bayesian techniques can be used to update trip rates developed by category analysis in a certain area before applying them in a different area. The methodology as well as example applications will be described hereafter for two cases:

1. New information obtained from small disaggregate sample coupled with the prior information based on the borrowed travel data from a similar urban area.

2. New information derived from subjective judgement coupled with the prior information based on the borrowed travel data from a similar urban area.

Small Disaggregate Sample

A trip rate table is borrowed from a certain area. This is the prior information. After processing the small sample, updating is done on a cell-by-cell basis, whereby each socio-economic combination (cell) is updated using sample information from the corresponding cell. Only cells that have at least 30 observations (in each of the prior and of the sample models) can be updated by this technique in order to meet the required assumptions.

Before the expressions for the updated parameters are presented, a brief description of the statistical procedure is necessary.

The following notation will be used:

For the cell being updated in the initial trip table (prior information), θ_1 : Average trip rate

n_1 : Number of observations

S_1^2 : Variance of the trip rates

For the corresponding cell (same socio-economic combination) developed from the small disaggregate sample,

θ_s : Average trip rate

n_s : Number of observations

S_s^2 : Variance of trip rates

The average trip rate for a certain cell is obtained by averaging individual household trip rates corresponding to that cell, and the average trip rates would follow a normal distribution regardless of the distribution of the individual household trip rates (by virtue of the Central Limit Theorem). The distribution of the average trip rate is referred to as the sampling distribution of the mean. This distribution is normal with an unknown mean. θ_s , the sample mean, is therefore normally distributed. It will be assumed that the variance σ_s^2 of this distribution is known, its value being $\sigma_s^2 = \frac{S_s^2}{n_s}$ (which is the unbiased estimator of the variance of the sampling distribution of the mean).

The prior distribution is the one derived from the initial trip table. Its mean is θ_1 and its variance is $\sigma_1^2 = \frac{S_1^2}{n_1}$ (unbiased estimators of the mean and variance of the sampling distribution of the mean for this case).

For the situation described above (both prior and sample distributions normal and variance assumed to be known), the updated distribution (posterior) of the average trip rate has the following parameters:

$$\text{mean } \theta_2 = \frac{1/\sigma_1^2}{(1/\sigma_1^2) + (1/\sigma_s^2)} \times \theta_1 + \frac{1/\sigma_s^2}{(1/\sigma_1^2) + (1/\sigma_s^2)} \times \theta_s$$

which becomes after substitution:

$$\theta_2 = \frac{n_1/S_1^2}{(n_1/S_1^2) + (n_s/S_s^2)} \times \theta_1 + \frac{n_s/S_s^2}{(n_1/S_1^2) + (n_s/S_s^2)} \times \theta_s \dots (\text{Eq. 1})$$

$$\text{Standard Deviation } \sigma_2 = \sqrt{\frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_s^2}}} \rightarrow \sigma_2 = \sqrt{\frac{S_1^2 \times S_s^2}{n_1 S_s^2 + n_s S_1^2}} \dots (\text{Eq.2})$$

It should be noted that σ_2 is the standard deviation of the sampling mean distribution, or of the average trip rate, as opposed to that of the individual household trip rate distribution (the former distribution being normal while the latter is unknown). The same comment could be made about θ_2 , even though in this case the difference is of a conceptual nature only, without practical implications since the means of the two distributions are but one and the same.

The above technique can be best understood through a detailed step-by-step example application. As a hypothetical example, it can be assumed that Lafayette's home based work productions trip rate table is to be used in Bloomington. A small disaggregate sample would be collected in Bloomington which would yield a trip rate table. The process of updating would then take place on a cell-by-cell basis. The cell corresponding to households of size 4 and owning 2 cars will serve as an illustration (c.f. Table 40):

Prior information:

$$\theta_1 = 2.28$$

$$n_1 = 236$$

$$S_1 = 1.642$$

Sample information (hypothetical):

$$\theta_s = 2.45$$

$$n_s = 31$$

$$S_s = 1.69$$

Posterior (Updated) information:

Using equation (1),

$$\begin{aligned}\theta_2 &= \frac{236/(1.652)^2}{[236/(1.652)^2] + [31/(1.69)^2]} \times 2.28 \\ &\quad + \frac{31/(1.69)^2}{[236/(1.652)^2] + [31/(1.69)^2]} \times 2.45 \\ &= 2.30\end{aligned}$$

and using equation (2),

$$\sigma_2 = \sqrt{\frac{(1.652)^2 \times (1.69)^2}{236 \times (1.69)^2 + 31 \times (1.652)^2}} = 0.101$$

Successive Updating

If another sample is collected at a later point in time, either in the same area or in another area for the purpose of updating further the information already available, a slightly different situation is created. The prior information in this case still consists of the sampling mean distribution, and it is still normally distributed. However, its parameters are different, and therefore computational formulas (1) and (2) will be different in form, even though they remain unchanged in content. The prior state of knowledge about that distribution at this state consists of the parameters θ_2 and σ_2 computed in the previous step. Therefore, the prior distribution is a normal distribution with mean θ_2 and variance σ_2^2 which are assumedly known. The new sample distribution will have the same notation as in the previous case (θ_s , n_s , S_s). No prior sample size is needed in order to

find the posterior distribution parameters. Let θ_3 and σ_3^2 be the mean and variance (respectively) of the posterior distribution. These are given by (56):

$$\text{mean } \theta_3 = \frac{1/\sigma_2^2}{(1/\sigma_2^2) + (n_s/S_s^2)} \times \theta_2 + \frac{1/\sigma_s^2}{(1/\sigma_2^2) + (1/\sigma_s^2)} \times \theta_s$$

which becomes after substitution:

$$\theta_3 = \frac{1/\sigma_2^2}{(1/\sigma_2^2) + (n_s/S_s^2)} \times \theta_2 + \frac{n_s/S_s^2}{(1/\sigma_2^2) + (n_s/S_s^2)} \times \theta_s \dots (\text{Eq.3})$$

$$\text{Std. Deviation } \sigma_3 = \sqrt{\frac{1}{\frac{1}{\sigma_2^2} + \frac{1}{\sigma_s^2}}} \rightarrow \sigma_3 = \sqrt{\frac{\sigma_2^2 \times S_s^2}{S_s^2 + n_s \sigma_2^2}} \dots (\text{Eq. 4})$$

To illustrate this last case, it can be assumed that a second sample was collected in Bloomington and yielded another trip table. This new information is to be used to improve the information already available (which itself is the result of the Lafayette trip table updated by the first Bloomington sample). The same cell-by-cell procedure will be adopted using the revised equations 3 and 4, as will be shown for the cell corresponding to households of size 4 and owning 2 cars:

Prior information(from previous updating):

$$\theta_2 = 2.30$$

$$\sigma_2 = 0.101$$

New sample information (hypothetical):

$$\theta_s = 2.55$$

$$n_s = 55$$

$$S_s = 1.72$$

Posterior (Updated) information:

Using Equation 3,

$$\begin{aligned}\theta_3 &= \frac{1/(0.101)^2}{[1/(0.101)^2] + [55/(1.72)^2]} \times 2.30 \\ &+ \frac{55/(1.72)^2}{[1/(0.101)^2] + [55/(1.72)^2]} \times 2.55 \\ &= 1.93 + 0.41 = 2.34\end{aligned}$$

Using Equation 4,

$$\sigma_3 = \sqrt{\frac{(0.101)^2 \times (1.72)^2}{(1.72)^2 + 55 \times (0.101)^2}} = 0.093$$

If need for further updating arises, the θ_3 and σ_3 just computed can serve as the prior distribution parameters, and the same equations (3 and 4) can be used again by substituting the appropriate values of the prior parameters as well as those of the new sample. The whole idea is to make maximum utilization of the information available prior and up to this point.

Subjective Judgement as a Source of New Information

Subjective judgement and evaluation of experienced planners familiar with the area of study can serve as a valuable source of new information in lieu of collecting a small disaggregate sample. Subjective judgement can also serve as the prior information if no prior information is available, while a small sample could provide the new information.

In order to use subjective information within a Bayesian framework, the primary task is to transform this subjective evaluation into mathematical values of distribution parameters. This can be done by having the planner specify a range within which the mean trip rate for a certain household category will fall, as well as the odds or chances that he is willing to give that this mean trip rate will fall within the specified range. This subjective information can be transformed into a mean and variance after a distribution for the planner's feelings about the parameter is assumed.

To illustrate this procedure, the same setting as in the previous example is assumed, except that this time the study area has decided not to collect a small disaggregate sample. Instead, an experienced planner familiar with the area is willing to give 4 to 5 odds that the mean home based work vehicle trip rate of households of size 4 and owning 2 cars lies between 2.1 and 2.7. A normal distribution would be an appropriate distribution to describe the planner's feelings about that parameter. The mean of this distribution is 2.4, and the variance can be calculated considering the fact that four fifths or 80 percent of the area under this normal curve falls between 2.1 and 2.7, therefore $1 - 0.8 = 0.2$ corresponds to the tail areas, i.e., $\alpha = \frac{0.2}{2} = 0.1$.

Using the normal distribution tail areas table (reproduced in any standard statistics reference), it turns out that an α of 0.1 corresponds to a Z-value of 1.28 (the details of these elementary steps will not be discussed). The standard deviation σ_s of this normal distribution can be evaluated from the following:

$$\frac{2.7 - \theta_s}{\sigma_s} = z$$

$$\frac{2.7 - 2.4}{\sigma_s} = 1.28$$

The above yields: $\sigma_s = \frac{2.7 - 2.4}{1.28} = 0.234$

All the necessary information is available for calculating the posterior distribution parameters. However, the formulas will look slightly different once the expressions for σ_1 and σ_s are substituted in the basic equations. In this case, σ_s is the standard deviation of the sampling mean distribution, and is available as such (unlike the case when the new information was supplied by a small sample, where $\sigma_s^2 = \frac{S_s^2}{n_s}$). The formulas for calculating the posterior distribution parameters are:

$$\text{mean } \theta_2 = \frac{1/\sigma_1^2}{(1/\sigma_1^2) + (1/\sigma_s^2)} \times \theta_1 + \frac{1/\sigma_s^2}{(1/\sigma_1^2) + (1/\sigma_s^2)} \times \theta_s$$

which becomes after substituting:

$$\theta_2 = \frac{n_1/S_1^2}{(n_1/S_1^2) + (1/\sigma_s^2)} \times \theta_1 + \frac{1/\sigma_s^2}{(n_1/S_1^2) + (1/\sigma_s^2)} \times \theta_s \dots (\text{Eq. 5})$$

$$\text{Std. Deviation } \sigma_2 = \sqrt{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_s^2}} \rightarrow \sigma_2 = \sqrt{\frac{S_1^2 \times \sigma_s^2}{n_1 \sigma_s^2 + S_1^2}} \dots (\text{Eq. 6})$$

The situation for this example is as follows

Prior Information (from Lafayette HBW trip table, cf. Table 40):

$$\theta_1 = 2.28$$

$$n_1 = 236$$

$$S_1 = 1.652$$

Sample information (subjective):

$$\theta_s = 2.4$$

$$\sigma_s = 0.234$$

Posterior (updated) information:

Using equations 5 and 6,

$$\begin{aligned}\theta_2 &= \frac{236/(1.652)^2}{[236/(1.652)^2] + [1/(0.234)^2]} \times 2.28 \\ &+ \frac{1/(0.234)^2}{[236/(1.652)^2] + [1/(0.234)^2]} \times 2.4 \\ &= 1.88 + 0.42 = 2.3\end{aligned}$$

$$\sigma_2 = \sqrt{\frac{(1.652)^2 \times (0.234)^2}{236 \times (0.234)^2 + (1.652)^2}} = 0.0977$$

The equations developed for all the cases described are summarized in Table 51.

Temporal updating is that where trip rates developed at an earlier point in time are updated to better reflect current or anticipated future conditions. Bayesian updating techniques can be used for this purpose as described in this chapter, therefore providing a very powerful tool for a continuing planning effort whereby new information (objective or subjective) obtained at a relatively low cost can be regularly used to update information available up to that point.

Conclusion

As described in this chapter, Bayesian statistics provide a very useful tool for updating trip generation rates and therefore improving

Table 51. Summary of Bayesian Updating Techniques for Trip Generation Rates

PRIOR		SAMPLE (NEW)		POSTERIOR DISTRIBUTION PARAMETERS	
Source	Parameters	Source	Parameters		
General zone	θ_1	General zone	θ_5	$\theta_2 = \left(\frac{1/\theta_1^2}{1/\theta_1^2 + 1/\theta_5^2} \right) \times \theta_1 + \left(\frac{1/\theta_5^2}{1/\theta_1^2 + 1/\theta_5^2} \right) \times \theta_5$	
	σ_1		σ_5	$\sigma_2 = \sqrt{\frac{1/\theta_1^2 \times \sigma_1^2}{1/\theta_1^2 + 1/\theta_5^2}}$	
From already developed trip table. One cell at a time	θ_1	Small Disaggregate Sample	θ_5	$\theta_2 = \left(\frac{n_1/\theta_1^2}{n_1/\theta_1^2 + n_5/\theta_5^2} \right) \times \theta_1 + \left(\frac{n_5/\theta_5^2}{n_1/\theta_1^2 + n_5/\theta_5^2} \right) \times \theta_5$	
	n_1	$\sigma_1^2 = \frac{S_1^2}{n_1}$	n_5		
	S_1		S_5	$\sigma_2 = \sqrt{\frac{S_1^2 \times S_5^2}{n_1^2 + n_5^2}}$	
From already updated trip tables, or from subjective information	θ_1	Small Disaggregate Sample	θ_5	$\theta_2 = \left(\frac{1/\theta_1^2}{1/\theta_1^2 + n_5/\theta_5^2} \right) \times \theta_1 + \left(\frac{n_5/\theta_5^2}{1/\theta_1^2 + n_5/\theta_5^2} \right) \times \theta_5$	
			n_5		
	σ_1		S_5	$\sigma_2 = \sqrt{\frac{1/\theta_1^2 \times S_5^2}{1/\theta_1^2 + n_5/\theta_5^2}}$	
From already developed trip table.	θ_1	Subjective Information	θ_5	$\theta_2 = \left(\frac{n_1/\theta_1^2}{n_1/\theta_1^2 + 1/\theta_5^2} \right) \times \theta_1 + \left(\frac{1/\theta_5^2}{n_1/\theta_1^2 + 1/\theta_5^2} \right) \times \theta_5$	
	n_1		n_5		
	S_1			$\sigma_2 = \sqrt{\frac{S_1^2 \times n_5^2}{n_1^2 + S_1^2}}$	

their accuracy in reflecting the trip making behavior of households. The cases described could be used for spatial as well as temporal updating, even though the major emphasis in this chapter was on the former.

Spatial updating is aimed at making trip rates borrowed from a similar urban area more descriptive of the households in the study area. As discussed earlier, this can be done using a small disaggregate sample collected in the study area or by taking advantage of the local expertise by quantifying the subjective judgement of local planners.

CHAPTER 8: CONCLUSION

This study has compared the trip frequency patterns in the small and medium sized urban areas (population range: 50,000 to 250,000) of Indiana where comprehensive travel surveys have been conducted. It was aimed at developing a framework for the transferability of trip frequency parameters at three levels of aggregation: areawide, zonal and household levels. Within this framework, the trip generation phase in the four-step Urban Transportation Planning process in small and medium sized urban areas can be performed synthetically, thus eliminating the need for costly internal home-interview O-D surveys. Procedures for synthetic trip generation will be recommended in this chapter based on the findings and accomplishments of this study which are summarized hereafter:

1. The distribution of trips by purpose is significantly different between the study areas.
2. Within an urban area, the distribution of trips by purpose varies between socio-economic groups as substantiated by the tests conducted on the household purpose distribution model developed using data from Evansville.
3. For small urban areas, differences in socio-economic characteristics constitute the major differentiating characteristics affecting trip frequency.

4. The socio-economic mix in the study areas, as indicated by the auto ownership distribution between households, is significantly different between these areas. Anderson and Muncie, however, do not seem to have significantly different auto ownership distributions, and this fact is reflected in the similarity of their areawide trip frequency rates.

5. Caution should be exercised when borrowing zonal trip generation equations for use in urban areas other than the one where they were developed.

6. Home Based Work zonal production equations have more potential for transferability than Home Based Other equations because of the nature of the work trip.

7. Home Based Other zonal production equations developed in the study areas led to erroneous predictions when used for trip generation in areas different from where they were developed.

8. Non Home Based zonal production equations have good potential for transferability and thus can be used successfully in synthetic applications.

9. Zonal trip attraction regression equations can be used to synthesize zonal attractions in areas different from the ones where they were developed. Non Home Based and Home Based Work synthetic equations predict more accurately than Home Based Other equations, even though the latter predict within reasonable limits. However, it is recommended that the synthesized attractions be balanced with total productions as obtained by more accurate synthetic techniques.

10. Disaggregate household trip generation rates have stronger theoretical justification for transferability because they are behavioral in nature. Empirical evidence to that effect was presented in this study based on the comparison of the cross-classified household trip rates from Lafayette and Evansville. This comparison showed that most of these rates were not significantly different for the two areas for all three trip purposes.

11. The number of levels needed within each variable in the cross-classification trip generation technique (when auto ownership and household size are the independent variables used) has been determined to be:

- Auto ownership: 0 cars, 1 car, 2 cars, 3 or more cars.
- Household size: 1, 2, 3, 4, 5 or more persons.

12. Bayesian statistics can be effectively used to update cross-classified household trip generation rates. Objective information derived from small disaggregate samples or subjective information based on expert engineering judgement can be used to update trip rates spatially as well as temporally.

Framework for the Transferability of Trip Frequency Parameters

A framework for the transferability of trip frequency parameters between small urban areas can be built based on the above conclusions. Instead of defining different classes of small urban areas within which parameters could be transferred at all levels of aggregation, this framework consists of the principles that should guide such transferring.

A fundamental premise adopted in this framework is that for small urban areas, socio-economic characteristics constitute the major differentiating elements affecting trip frequency. Areawide frequency parameters, such as areawide average trip rates per household or per capita, are therefore transferable between areas only when these areas have similar socio-economic characteristics. Of course, for rough areawide estimates, "typical" rates can be used. However, more accurate analyses require more confidence in the parameters used.

In order to be transferable between small urban areas, trip generation models and relationships should relate travel to its basic determinants. Zonal aggregate regression equations tend to mask the causal aspects of the assumed relationships, besides the fact that they are tied to a specific zonal scheme. Therefore, extreme caution should be exercised if such relationships are to be transferred.

On the other hand, disaggregate household models, because they are behavioral in nature, have the highest potential as well as the strongest theoretical justification for transferability. Household trip generation rates are therefore more readily transferable between small and medium sized urban areas.

Within this framework, trip generation analysis can be synthetically conducted for small urban areas. A vast array of techniques and approaches are available for this purpose, and most of them have been discussed in this study. A recommended procedure for synthetic trip generation analysis is outlined hereafter.

Recommended Procedure for Synthetic Trip Generation

Analysis in Small Urban Areas

As explained in earlier chapters, this procedure is limited to internal vehicle trips.

Before detailed trip generation analysis is undertaken, rough estimates can be obtained using the rates reported in Chapters 3 and 4.

The following steps could next be followed:

a. Establishing Control Totals

This could be done using the disaggregate trip rate tables as explained in Chapter 6, whereby areawide control totals can be obtained by multiplying the appropriate rate by the areawide number of households in each category.

Control totals are obtained for each purpose separately. It should be noted however that the Non-Home Based trip rates reported in Chapter 6 do not include truck trips. The missing information can be supplied either by using the areawide vehicle trip rates reported in Chapter 4, or by using local survey information, or by multiplying the number obtained from the trip rates by a correction factor between 1.4 and 1.6 depending on the area being studied (this range is based on figures from the study areas).

b. Trip Productions

Zonal values of the trip productions are needed. These can be obtained in the following way:

- HBW and HBO: using cross-classification tables and multiplying by the zonal number of households of each category. HBW productions can alternatively be determined using the equation recommended for that purpose in Chapter 5.
- NHB: using synthetic equations as recommended in Chapter 5.

The above zonal productions should be adjusted to agree with the control totals established earlier.

c. Trip Attractions

As described in Chapter 5, regression equations can be used to synthesize zonal attractions provided that the total attractions are balanced by total productions for each purpose. In addition, special analyses might be needed for special generators unique to the study area.

Alternatively, attraction rates derived from national data (41, 42) can be used as described in Chapter 5, again with the condition that the total attractions are constrained to equal the total productions.

Finally, the whole process would have to be "tuned" so that the volumes assigned to the network (after the trip distribution and trip assignment steps are executed) compare favorably with the observed volume counts.

Recommendations for Further Research

Synthetic trip generation for small urban areas is an activity that is in need of continuous research associated with specific applications to study areas. There is a definite lack of documented

experience in this activity which is not being carried out on a large scale. The body of knowledge in this field can significantly benefit from applications to actual urban areas.

Some specific research needs are:

- Further substantiation of the transferability of disaggregate household trip rates by conducting further comparisons similar to the one described in Chapter 6.
- Review and refinement of the existing techniques used for forecasting the independent variables needed as inputs for trip generation in the future year.
- Investigation of external trips and development of a procedure for synthesizing them for small urban areas.
- Application of Bayesian updating to a specific case study and evaluation of the improvement due to such updating.

The recommendations and findings of this study serve to drastically reduce the time and cost required by the transportation planning process in small and medium-sized areas. These savings in resources can thus be directed towards other pressing and current transportation issues and problems.

LIST OF REFERENCES

LIST OF REFERENCES

1. Heanue, K. E., "Changing Emphasis in Urban Transportation Planning", presented at the 56th Annual Meeting of the Transportation Research Board, Washington, D. C., 1977.
2. Levinson, H. S., "Characteristics of Urban Transportation Demand: A New Data Bank", presented at the 57th Annual Meeting of the Transportation Research Board, Washington, D. C., January 1978.
3. Sosslau, A. B.; Hassam, A. B. and Carter, M. M., "Manual Techniques and Transferable Parameters for Urban Transportation Planning", presented at the 57th Annual Meeting of the Transportation Research Board, Washington, D. C., January 1978.
4. Bolsan, R., "New Rules for Judging Analytical Techniques in Urban Planning", in ASPO, Analytical Techniques, 1970, pp. 75-84.
5. Roberts, P. O., "Model System for Urban Transportation Planning: Where Do We Go From Here?", Highway Research Record 309, 1970, pp. 34-44.
6. Lee, D. B., "The Future of Models in Transportation", presented at the 1971 Annual Meeting of the American Institute of Planners, San Francisco, October 1971.
7. Lee, D. B., "Requiem for Large-Scale Models", Journal of the American Institute of Planners, Vol. 39, No. 3, May 1973, pp. 163-178.
3. Kassoﬀ, H. and Deutschman, H. D., "Trip Generation: A Critical Appraisal", Highway Research Record 297, 1969, pp. 15-30.
9. Hartgen, D. and Tanner, G., "Behavioral Models of Mode Choice", New York State DOT, March 1970.
10. Brand, D. "Theory and Methods in Land Use and Travel Demand Forecasting", paper presented at the Highway Research Board Meeting, Washington, D. C., January 1972.
11. Lave, C. A., "A Behavioral Approach to Modal Split Forecasting", Transportation Research, Vol. 3, No. 4, December 1969, pp. 463-480.
12. McGillivray, R. G., "Some Problems with Urban Transportation Planning Models", Traffic Quarterly, Vol. 26, No. 4, October 1972, pp. 547-559.

13. Wickstrom, G. V. and Sosslau, A. B., "Travel Estimation Procedures for Quick Response to Urban Policy Issues", NCHRP Interim Report for Project 8-12, Washington, D. C., April 1976.
14. Federal Highway Administration, "Transportation Planning in Small Urban Areas: The Appropriate Level of Effort", February 1977.
15. Organization for Economic Cooperation and Development Road Research Group, "Urban Traffic Models: Possibilities for Simplification", August 1974.
16. Grecco, W. L. et al., "Transportation Planning for Small Urban Areas", NCHRP Report 167, Transportation Research Board, 1976.
17. Coomer, B. and Corradino, J., "Trip Generation Distribution in a Small Urban Area -- An Efficiency Analysis", Traffic Engineering, Vol. 43, June 1973, pp. 60-67.
18. Chan, Y., "Review and Compilation of Demand Forecasting Experiences: An Aggregation of Estimation Procedures", Pennsylvania Transportation Institute, June 1977.
19. Bates, J., "Development and Testing of Synthetic Generation and Distribution Models for Urban Transportation Studies", State Highway Department of Georgia, Atlanta, 1971.
20. Kristoffersen, S. and Wilson, E. M., "Trip Generation Synthesis for Small and Medium Sized Cities", presented at the 56th Annual Meeting of the Transportation Research Board, January 1977.
21. Jefferies, W. R. and Carter, E. C., "Simplified Techniques for Developing Transportation Plans", Highway Research Record 240, 1968, pp. 60-87.
22. Federal Highway Administration, "Trip Generation Analysis", August 1975.
23. Khasnabis, S. and Poole, M., "Synthesizing Travel Patterns for a Small Urban Area", Traffic Engineering, Vol. 45, No. 8, August 1975.
24. Martinson, D. R., Chatterjee, A. and Sinha, K. C., "Trip Generation Analysis for Regional Studies", TE Journal of the ASCE, Vol. 103, November 1977.
25. Wooton and Pick, "A Model for Trips Generated by Households", Journal of Transport Economics and Policy, 1, 1967, pp. 137-153.
26. Neter, J. and Wasserman, W., Applied Linear Statistical Models, Homewood, Illinois: Richard D. Irwin, Inc., 1974.

27. Vogt, Sage and Pflum Consultants, "Traffic Simulation Model Procedures Report for Columbus-Bartholomew County Transportation Study", June 1976.
28. COMSIS Corporation, "Quick-Response Urban Travel Estimation Manual Techniques and Transferable Parameters -- A Users' Guide", Final Report, NCHRP project 8-12A, November 1977.
29. Gilbert, G. and Dajani, J., "Energy, Urban Form and Transportation Policy", Transportation Research, Vol. 8, No. 415, 1974, pp. 267-276.
30. Jamieson, G. B.; Mackay, W. K. and Latchford, J. C. R., "Transportation and Land Use Structures", Urban Studies, Vol. 4, No. 3, November 1967, pp. 201-217.
31. Edwards, J. L. and Schofer, J. L., "Relationships Between Transportation Energy Consumption and Urban Structure: Results of Simulation Studies", Transportation Research Record 599, 1975.
32. Hemmens, G. C., "Experiments in Urban Form and Structure", Highway Research Record 207, 1967, pp. 32-41.
33. Schneider, J. I. and Beck, J. D., "Reducing the Travel Requirements of the American City: An Investigation of Alternative Urban Spatial Structures", Transportation Research Record 499, 1974, pp. 12-33.
34. Oi and Shuldiner, An Analysis of Urban Travel Demand, Northwestern University Press, 1962.
35. Draper and Smith, Applied Regression Analysis, Wiley and Sons, 1966.
36. Hutchinson, B. G., Principles of Urban Transport Systems Planning, Scripta Book Co., Washington, D. C., 1974.
37. Dickey, J. W., Metropolitan Transportation Planning, Scripta Book Co., Washington, D. C., 1975.
38. Paquette, Ashford and Wright, Transportation Engineering: Planning and Design, The Ronald Press Co., New York, 1972.
39. Federal Highway Administration, "Guidelines for Trip Generation Analysis", June 1967.
40. Sexton, B. J., "Kentucky's Transportation Planning Program for Small Urban Areas", Traffic Quarterly, October 1969, pp. 561-571.
41. Institute of Transportation Engineers, "Trip Generation: An Informational Reprot", Arlington, Virginia, 1976.
42. Arizona Department of Transportation, "T.G.I.F.: Trip Generation Intensity Factors", Reprinted by the U. S. DOT/FHWA, June 1977.

43. McCarthy, G. M., "Multiple Regression Analysis of Household Trip Generation--A Critique", Highway Research Record 297, 1969, pp. 31-43.
44. Kannel, E. J. and Heathington, K. W., "Temporal Stability of Trip Generation Relations", Highway Research Record 472, 1973, pp. 17-27.
45. Dobson, R. and McGarvey, W.E., "An Empirical Comparison of Disaggregate Category and Regression Trip Generation Techniques", Transportation, 6, 1977, pp.287-307.
46. Chatterjee, A. and Khasnabis, S., "Category Models -- A Case for Factorial Analysis", Traffic Engineering, Vol. 44, No. 1, October 1973, pp. 29-33.
47. Kirkpatrick, E. G., Introductory Statistics and Probability for Engineering, Science, and Technology. Printice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.
48. Ashford, N. and Holloway, M., "Time Stability of Zonal Trip Production Models", TE Journal of the ASCE, November 1972, pp. 799-806.
49. Martinson, D. R., "A Practical Approach to Trip Generation Analysis for a Multi-County Region", unpublished Masters Thesis, Marquette University, Milwaukee, Wisconsin, July 1974.
50. Ben-Akiva, M. E. and Atherton, T. J., "Transferability and Updating of Disaggregate Travel Demand Models", TRB Record 610, 1976, pp. 12-18.
51. Raiffa, H. and Schlaifer, R., Applied Statistical Decision Theory, Division of Research, Graduate School of Business Administration, Harvard University, Boston, 1961.
52. Isibor, E. I., "Modeling the Impact of Highway Improvements on the Value of Adjacent Land Parcels", Joint Highway Research Project - Purdue University, Report 37, 1969.
53. Sinha, K. C., "Statistical Decision in Forecasting Planning Data", Transportation Engineering Journal, A.S.C.E., Vol. 98, No. TE4, Proc. Paper 9366, November 1972, pp. 865-880.
54. Chatterjee, A. and Sinha, K. C., "Mode Choice Estimation for Small Urban Areas", Transportation Engineering Journal, A.S.C.E., Vol. 101, No. TE2, Proc. Paper 11303, May 1975, pp. 265-278.
55. Saraf C. et al., "A Bayesian Approach Applied to the Prediction of Pavement Distress", Paper presented at the 57th Annual Meeting of the Transportation Research Board, Washington, D. C., January 1978.
56. DeGroot, M. H., Optimal Statistical Decisions, McGraw-Hill Book Company, 1970.

General References

Baerwald, John E., ed., Transportation and Traffic Engineering Handbook, Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1976.

Baker, R. F., ed., Handbook of Highway Engineering, New York, N. Y.: Van Nostrand Reinhold Company, 1975.

Burr, L. W., Applied Statistical Methods, New York, N. Y.: Academic Press, 1974.

Krueckeberg, D. A. and Silvers, A. L., Urban Planning Analysis: Methods and Models, New York, N. Y.: John Wiley and Sons, Inc., 1974.

APPENDIX

APPENDIX

Contingency Tables*

Contingency tables constitute an important class of problems to which the Chi-square test can be applied. These in general are in the form of a rectangular table of frequencies. Two factors are associated with a contingency table: a row factor A and a column factor B. Usually, the hypothesis being tested (null hypothesis) is that the two factors are independent. This is often called "independence of classification". The alternative hypothesis is that the two factors do not act independently; that is, that they interact. A Chi-square test can be used for that purpose, whereby expected cell frequencies are computed and compared to the observed cell frequencies. The general setup of a contingency table is described hereafter.

General Setup of a Contingency Table

Table A1 gives a general setup for an $a \cdot b$ contingency table. The observed frequencies are

$$f_{ij}, i = 1, \dots, a; j = 1, \dots, b.$$

The expected cell frequencies F_{ij} are given by:

$$F_{ij} = (n_{i.})(n_{.j})/n.$$

where

$$F_{ij} = \text{expected frequency in cell } i, j$$

$$n_{i.} = \text{marginal total in row } i$$

Table A1. General Contingency Table, a.b

Factor A	Factor B				Total
	1	2 . . . j . . . b			
1	f_{11}	f_{12}	f_{1j}	f_{1b}	$n_{1.}$
2	f_{21}	f_{22}	f_{2j}	f_{2b}	$n_{2.}$
.					
.					
.					
i	f_{i1}	f_{i2}	f_{ij}	f_{ib}	$n_{i.}$
.					
.					
.					
a	f_{a1}	f_{a2}	f_{aj}	f_{ab}	$n_{a.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.j}$	$n_{.b}$	n

$n.j$ = marginal total in column j

$n.$ = grand total,

as shown in Table A1.

It should be noted that:

$$\sum_{i=1}^a F_{ij} = n.j ,$$

$$\sum_{j=1}^b F_{ij} = n_i . ,$$

$$\sum \sum F_{ij} = n.$$

The expected cell frequencies (F_{ij} 's) are compared to the observed cell frequencies (f_{ij} 's) by:

$$\chi^2_{data} = \sum_{i=1}^a \sum_{j=1}^b (f_{ij} - F_{ij})^2 / F_{ij}$$

This carries degrees of freedom

$$v = (a - 1) (b - 1)$$

This test statistic would follow the Chi-Square Distribution if the two factors were independent. For this reason, it is compared to the critical value for the Chi-Square Distribution at the desired significance level. These critical values are found in standard tables reported in any probability and statistics textbook. If the computed statistic (χ^2_{data}) is equal to or less than the critical value $\chi^2_{v,\alpha}$ (α being the significance level), the null hypothesis is accepted; otherwise, it is rejected.

*Reference: Burr, I.W., Applied Statistical Methods, Academic Press, 1974.

COVER DESIGN BY ALDO GIORGINI